Journal of Mathematical Physics

Nonlinear PDEs approach to statistical mechanics of dense associative memories

Cite as: J. Math. Phys. **63**, 103304 (2022); https://doi.org/10.1063/5.0095411
Submitted: 10 April 2022 • Accepted: 21 September 2022 • Published Online: 27 October 2022

Elena Agliari, Alberto Fachechi and Chiara Marullo

Journal of Mathematical Physics | ARTICLE | scitation.org/journal/jmp

Nonlinear PDEs approach to statistical mechanics of dense associative memories

Cite as: J. Math. Phys. 63, 103304 (2022); doi: 10.1063/5.0095411
Submitted: 10 April 2022 • Accepted: 21 September 2022 •
Published Online: 27 October 2022

View Online | Export Citation | CrossMark

Elena Agliari,[1,2,a)] Alberto Fachechi,[1,2] and Chiara Marullo[1,2]

AFFILIATIONS

[1] Dipartimento di Matematica "Guido Castelnuovo," Sapienza Università di Roma, Roma, Italy
[2] GNFM-INdAM, Gruppo Nazionale di Fisica Matematica, Istituto Nazionale di Alta Matematica, Lecce, Italy

[a)] **Author to whom correspondence should be addressed:** agliari@mat.uniroma1.it

ABSTRACT

Dense associative memories (DAMs) are widely used models in artificial intelligence for pattern recognition tasks; computationally, they have been proven to be robust against adversarial inputs and, theoretically, leveraging their analogy with spin-glass systems, they are usually treated by means of statistical-mechanics tools. Here, we develop analytical methods, based on nonlinear partial differential equations, to investigate their functioning. In particular, we prove differential identities involving DAM's partition function and macroscopic observables useful for a qualitative and quantitative analysis of the system. These results allow for a deeper comprehension of the mechanisms underlying DAMs and provide interdisciplinary tools for their study.

I. INTRODUCTION

Artificial intelligence (AI) is rapidly changing the face of our society thanks to its impressive abilities in accomplishing complex tasks and extracting information from nontrivially structured, high-dimensional datasets. The successful applications of modern AI range from hard sciences and technology to more practical scenarios (e.g., medical sciences, economics and finance, and many daily tasks). Its success is primarily due to the ascent of deep learning,[1,2] a set of semi-heuristic techniques consisting of several minimal building blocks stacked together in complex architectures with extremely high learning performances. Despite its success, a rigorous theoretical framework guiding the development of such machine-learning architectures is still lacking. In this context, statistical mechanics of complex systems offers ideal tools to study neural-network models from a more theoretical (and rigorous) point of view, thus drawing a feasible path that makes AI less empirical and more explainable.

In statistical mechanics, a primary classification of physical systems is the following. On the one hand, we have simple systems, which are essentially characterized by the fact that the number of equilibrium configurations does not depend on the system size N. A paradigmatic (mean-field) realization of this situation is the Curie–Weiss (CW) model in which all the spins σ_i, $i = 1, \ldots, N$, making up the system interact pairwise by a constant, positive (i.e., ferromagnetic) coupling J. Below the critical temperature, in fact, the system exhibits ordered collective behaviors, and the equilibrium configurations of the system are characterized by only two possible values of the global magnetization $m(\boldsymbol{\sigma}) := \frac{1}{N}\sum_{i=1}^{N} \sigma_i$ (which are further related by a spin-flip symmetry $\sigma_i \to -\sigma_i$ for each $i = 1, \ldots, N$). On the other hand, we have complex systems in which the number of equilibrium configurations increases according to an appropriate function of the system size N due to the presence of frustrated interactions.[3] The prototypical example of mean-field spin-glass is the Sherrington–Kirkpatrick (SK) model[4] in which the interaction strengths between the spin pairs are i.i.d. Gaussian variables. With respect to simple systems, spin-glass models exhibit a richer physical and mathematical structure, as shown by the presence of the spontaneous replica-symmetry breaking and an infinite number of phase transitions (e.g., see Refs. 5–12) as well as the ultrametric organization of pure states (e.g., see Refs. 13–15). Statistical mechanics of spin-glasses has acquired a prominent role during the last decades due to its ability to describe the equilibrium dynamics of several paradigmatic models for AI, in particular thanks to the work of Amit, Gutfreund, and Sompolinsky[16] on associative neural networks. For our concerns, the relevant ones are the Hopfield model[17,18] and its p-spin extensions, the Dense Associative Memories (DAMs),[19–21] exhibiting features that are peculiar to both ferromagnetic (simple) and spin-glass (complex) systems. In these models, the interactions between the spins are designed in order to store K "information patterns" denoted by $\{\boldsymbol{\xi}^\mu\}_{\mu=1,\dots,K}$, where $\boldsymbol{\xi}^\mu = (\xi_1^\mu, \xi_2^\mu, \dots, \xi_N^\mu) \in \{-1,+1\}^N$ and ξ_i^μ is a Rademacher random variable for any $i = 1,\dots,N$ and $\mu = 1,\dots,K$; the μ-th pattern is said to be stored if the configuration $\boldsymbol{\sigma} = \boldsymbol{\xi}^\mu$ is an equilibrium state and the relaxation to this configuration, starting from a relatively close one (i.e., a corrupted version of $\boldsymbol{\xi}^\mu$), is interpreted as the retrieval of that pattern. The Hamilton function (or the energy in physics jargon) of these systems can be expressed as

$$H_{N,p}(\boldsymbol{\sigma}) \propto -\sum_{\mu=1}^{K} (m_\mu(\boldsymbol{\sigma}))^p,$$

where p is the interaction order (for the Hopfield model $p = 2$, while $p > 2$ for the DAMs) and $m_\mu(\boldsymbol{\sigma}) := \frac{1}{N}\sum_{i=1}^N \xi_i^\mu \sigma_i$ is the so-called Mattis magnetization measuring the retrieval of the μ-th pattern. It has been shown that the number of storable patterns scales, at most as a function of the system size, more precisely $K < \alpha_c(p)N^{p-1}$, where $\alpha_c(p) \in \mathbb{R}$, depends on the interaction order p and is referred to as critical storage capacity.[22,23] By a statistical-mechanics investigation of these models, one can highlight the macroscopic observables (order parameters) useful to describe the overall behavior of the system, i.e., to assess whether it exhibits retrieval capabilities, and the natural control parameters whose tuning can qualitatively change the system behavior; such knowledge can then be summarized in *phase diagrams*.

Regarding the methods, we can rely on a wide set of techniques for analyzing the equilibrium dynamics of complex systems and, in particular, to solve for their free energy. Historically, the first method (which was applied to the SK model and the Hopfield model[16,24]) is the *replica trick*, which—despite being straightforward and effective—is semi-heuristic and suffers from delicate points, see for example Ref. 25. Alternatively, rigorous approaches have also been developed and, among these, the relevant one for our concerns is Guerra's interpolating framework. In this case, we can take advantage of rigorous mathematical methods by applying sum rules[26] or by mapping the relevant quantities (the free energy or the model order parameters) of the statistical setting to the solutions of PDE systems. Indeed, differential equations involving the partition functions (or related quantities) of thermodynamic models have been extensively investigated in the literature, see for example Refs. 27–37. In particular, they allow us to express the equation of state governing the equilibrium dynamics of the system in terms of solutions of nonlinear differential equations and to describe phase transition phenomena as a development of shock waves, thus linking critical behaviors to gradient catastrophe theory.[38–41] In a recent study,[36] a direct connection between the thermodynamics of ferromagnetic models with interactions of order p and the equations of the Burgers hierarchy was established by linking the solution of the latter to the equilibrium solution of the order parameter of the former (i.e., the global magnetization m). In the present paper, we extend these results to complex models, in particular, to the Hopfield model and the DAMs.

The paper is organized as follows: In Sec. II, we introduce the relevant tools for our investigations, in particular Guerra's interpolating scheme for the PDE duality. In Sec. III, as a warm up, we review some basic results about the p-spin ferromagnetic models. In Sec. IV, we extend our results to the Derrida model (constituting the p-spin extension of the SK spin glass).[42] In Sec. V, we merge our results in a unified methodology for dealing with the DAMs, especially in the so-called high-storage limit, and re-derive the self-consistency equations for the order parameters by means of PDE technology.

II. GENERALITIES AND NOTATION

In this section, we present the thermodynamic objects we aim to study. We start with a system made up of N spins whose configurations $\boldsymbol{\sigma} \in \Sigma_N \equiv \{-1,+1\}^N$ are the nodes of a hypercube and that interact via a suitable tensor $\boldsymbol{J}$ of order p. The Hamilton functions of the system we will consider in this paper are of the form

$$H_{N,p,\boldsymbol{J}}(\boldsymbol{\sigma}) = -\frac{1}{D_{p,N,J}} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1,i_2,\dots,i_p}\sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_p}, \tag{2.1}$$

where $D_{p,N,J}$ is a normalization factor ensuring the linear extensivity of the energy with the system size. Once the Hamiltonian is fixed, we introduce the partition function in the usual Boltzmann–Gibbs form. Thus, given $\beta \in \mathbb{R}_+$ the level of thermal noise of the system, the partition function is defined as

$$Z_{N,p,\boldsymbol{J}}(\beta) := \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp[-\beta H_{N,p,\boldsymbol{J}}(\boldsymbol{\sigma})]. \tag{2.2}$$

As is standard in statistical mechanics, it is convenient to compute intensive quantities that are well-defined in the thermodynamic limit $N \to \infty$. Since the partition function is a sum of 2^N contribution, it is sufficient to take the intensive logarithm of the partition function, i.e.,

$$A_{N,p,\boldsymbol{J}}(\beta) := \frac{1}{N}\log Z_{N,p,\boldsymbol{J}}(\beta),$$

which is the intensive statistical pressure (which, apart from a factor $-\beta$, is the usual free energy) of the system. When dealing with spin-glass systems, the coupling tensor $\boldsymbol{J}$ is a multidimensional random variable, thus the partition function defines a random measure on the configuration space. For good enough probability distributions of the coupling matrix, the intensive logarithm of the partition function is expected to converge to its expectation value in the thermodynamic limit $N \to \infty$ by virtue of self-averaging theorems,[43,44] so it is natural to consider the quenched intensive pressure associated with the partition function (2.2), which is defined as

$$A_{N,p}(\beta) := \frac{1}{N}\mathbb{E}_{\boldsymbol{J}} \log Z_{N,p,\boldsymbol{J}}(\beta), \tag{2.3}$$

where $\mathbb{E}_{\boldsymbol{J}}$ denotes the average over the quenched disorder $\boldsymbol{J}$ (we stress that, in this case, the free energy does not depend any longer on the coupling matrix because of the average operation).

Rather than working with the quantity (2.2), we will use as a fundamental object Guerra's interpolating partition function and its associated interpolating intensive pressure. For instance, for spin-glass systems, we would have

$$\begin{aligned} Z_{N,p,\boldsymbol{J}}(t,\boldsymbol{x}) &:= \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp[-H_{N,p,\boldsymbol{J}}(t,\boldsymbol{x})],\\ A_{N,p}(t,\boldsymbol{x}) &:= \frac{1}{N}\mathbb{E}_{\boldsymbol{J}} \log Z_{N,p,\boldsymbol{J}}(t,\boldsymbol{x}), \end{aligned} \tag{2.4}$$

where $H_{N,p,\boldsymbol{J}}(t,\boldsymbol{x})$ denotes the interpolating Hamiltonian satisfying the properties such that (1) at $\boldsymbol{x} = 0$ and $t \neq 0$, it recovers the Hamiltonian (2.1) times β, and (2) at $t = 0$ and $\boldsymbol{x} \neq \boldsymbol{0}$, it corresponds to an exactly-solvable one-body system—a system where spins interact only with an external field that has to be set *a posteriori*. The interpolating parameters t and $\boldsymbol{x}$ are interpreted, in a mechanical analogy, as spacetime coordinates with suitable dimensionality.

The interpolating structure (2.4) implies a generalized measure, whose related Boltzmann factor is

$$B_{N,p,\boldsymbol{J}}(t,\boldsymbol{x}) := \exp[-H_{N,p,\boldsymbol{J}}(t,\boldsymbol{x})]. \tag{2.5}$$

Thus, for an arbitrary observable $O(\boldsymbol{\sigma})$ in the configuration space Σ_N, we can introduce the Boltzmann average induced by the partition function (2.4) as

$$\omega_{t,\boldsymbol{x}}(O) := \frac{1}{Z_{N,p,\boldsymbol{J}}(t,\boldsymbol{x})} \sum_{\boldsymbol{\sigma}\in\Sigma_N} O(\boldsymbol{\sigma}) B_{N,p,\boldsymbol{J}}(t,\boldsymbol{x}). \tag{2.6}$$

Usually, in spin-glass systems, the quenched average is performed after taking the Boltzmann expectation values on the s-replicated space $\Sigma_N^{(s)} = (\Sigma_N)^{\otimes s} \equiv \{-1,+1\}^{sN}$, which is naturally endowed with a random Gibbs measure corresponding to the partition function $Z^{(s)}_{N,p,\boldsymbol{J}}(t,\boldsymbol{x}) = Z_{N,p,\boldsymbol{J}}(t,\boldsymbol{x})^s$. Given a function $O : \Sigma_N^{(s)} \to \mathbb{R}$, the Boltzmann average in the s-replicated space is straightforwardly defined as

$$\Omega^{(s)}_{t,\boldsymbol{x}}(O) := \frac{1}{Z^{(s)}_{N,p,\boldsymbol{J}}(t,\boldsymbol{x})} \sum_{\underline{\boldsymbol{\sigma}}\in\Sigma_N^{(s)}} O(\underline{\boldsymbol{\sigma}}) B^{(s)}_{N,p,\boldsymbol{J}}(t,\boldsymbol{x}),$$

where $\underline{\boldsymbol{\sigma}} \in \Sigma_N^{(s)}$ is the global configuration of the replicated system and $B^{(s)}_{N,p,\boldsymbol{J}}(t,\boldsymbol{x})$ is the Boltzmann factor associated with the s-replicated partition function. Of course, in spin-glass theory, the relevant quantities are the *quenched* expectation values, which are defined as

$$\langle O\rangle_{t,\boldsymbol{x}} := \mathbb{E}_{\boldsymbol{J}}\Omega^{(s)}_{t,\boldsymbol{x}}(O). \tag{2.7}$$

For the sake of simplicity, we drop the index s from the quenched averages, as it would be clear from the context.

With all these definitions in mind, we are then able to find the link between the resolution of the statistical mechanics of a given spin-like model and a specific PDE problem in the fictitious space $(t,\boldsymbol{x})$. Before concluding this section, it is worth recalling that here we will work under the replica-symmetry (RS) assumption, meaning that we assume the self-averaging property for any order parameter X, i.e., the fluctuations around their expectation values vanish in the thermodynamic limit. In distributional sense, this corresponds to

$$\lim_{N\to\infty} \mathcal{P}_{t,\boldsymbol{x}}(X) = \delta(X - \bar{X}), \tag{2.8}$$

where $\bar{X} = \langle X\rangle_{t,\boldsymbol{x}}$ is the expectation value with respect to the interpolating measure $\mathcal{P}_{t,\boldsymbol{x}}(X)$. Typically, for simple systems, this assumption is correct. However, for complex systems, this is not always the case; for instance, in spin-glasses, the RS is broken at low temperature.[3] When dealing with neural-network models, RS constitutes a standard working assumption as it usually applies (at least) in a limited region of the parameter space, while elsewhere it yields only small quantitative discrepancies with respect to the exact solution.[45,46] The latter, accounting for RSB phenomena, can be obtained by iteratively perturbing the RS interpolation scheme (e.g., see Refs. 37, 47, and 48); thus, our results find direct application on the practical side and provide the starting point for further refinements on the theoretical side.

The models that shall be addressed in the next sections display Hamiltonian functions (2.1) that only differ in the choice of the tensor $\boldsymbol{J}$: We will start with the simplest case in which the interaction strength between the spins is homogeneous (i.e., $J_{i_1,i_2,\ldots,i_p} = J \in \mathbb{R}^+$); then, we will move to a more complex system, the Derrida model, where the interaction strength is randomly drawn (i.e., $J_{i_1,i_2,\ldots,i_p} \sim \mathcal{N}(0,1)$), finally ending with DAMs in which $\boldsymbol{J}$ is a tensorial generalization of the standard Hebbian storing rule, where the K patterns $\{\boldsymbol{\xi}^\mu\}_{\mu=1,\ldots,K}$ are allocated (i.e., $J_{i_1,i_2,\ldots,i_p} = N^{1-p}\sum_{\mu=1}^K \xi_{i_1}^\mu \ldots \xi_{i_p}^\mu$).

Our main results consist in proving that, *i*, the expectation value of the order parameter of the Derrida model is nothing but the solution of the p-th Burgers hierarchy, in the inviscid limit and under appropriate initial conditions, *ii.* for DAMs networks, the expectation values of the order parameters satisfy a system of PDEs with Burgers-hierarchy-like structure.

III. p-SPIN FERROMAGNETIC MODELS: HOW TO DEAL WITH SIMPLE SYSTEMS

The present section is a compendium of the results reported in Ref. 36, so we refer to that work for a detailed derivation. In p-spin ferromagnets, the interaction between spins is fixed by the requirement $J_{i_1,i_2,\ldots,i_p} = J$ for each $i_1,\ldots,i_p = 1,\ldots,N$ and $J > 0$; without loss of generality, one can set $J = 1$, since it corresponds to a rescaling of the thermal noise. Thus, the Hamilton function of the model simply reads as

$$H_{N,p}(\boldsymbol{\sigma}) := -\frac{1}{N^{p-1}} \sum_{i_1,\ldots,i_p=1}^{N} \sigma_{i_1}\ldots\sigma_{i_p} = -N(m(\boldsymbol{\sigma}))^p, \tag{3.1}$$

with

$$m(\boldsymbol{\sigma}) := \frac{1}{N}\sum_i \sigma_i$$

being the global magnetization of the system. By following the same lines of Ref. 36, Guerra's interpolating partition function reads

$$Z_{N,p}(t,x) = \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp(-H_{N,p}(t,x)), \tag{3.2}$$

$$H_{N,p}(t,x) = tNm(\boldsymbol{\sigma})^p - Nxm(\boldsymbol{\sigma}), \tag{3.3}$$

where $(t,x) \in \mathbb{R}^2$. The starting point is to notice that the interpolating statistical pressure associated with the partition function (3.2) has spacetime derivatives given by

$$\partial_t A_{N,p}(t,x) = -\omega_{t,x}(m(\boldsymbol{\sigma})^p), \tag{3.4}$$

$$\partial_x A_{N,p}(t,x) = \omega_{t,x}(m(\boldsymbol{\sigma})). \tag{3.5}$$

The expectation value of monomials of the global magnetization satisfies the following relation:[36]

$$\partial_x \omega_{t,x}(m(\boldsymbol{\sigma})^s) = N(\omega_{t,x}(m(\boldsymbol{\sigma})^{s+1}) - \omega_{t,x}(m(\boldsymbol{\sigma})^s)\omega_{t,x}(m(\boldsymbol{\sigma}))). \tag{3.6}$$

This means that we can act on the expectation value $\omega_{t,x}(m(\boldsymbol{\sigma}))$ to generate higher momenta. In particular, calling $u(t,x) = \omega_{t,x}(m(\boldsymbol{\sigma}))$ and setting $s = p-1$, we directly get the Burgers hierarchy,

$$\partial_t u(t,x) + \partial_x\left(\frac{1}{N}\partial_x + u(t,x)\right)^{p-1} u(t,x) = 0. \tag{3.7}$$

This duality also allows us to analyze the thermodynamic limit, corresponding to the inviscid scenario for the Burgers hierarchy. Indeed, posing $\bar{u}(t,x) = \lim_{N\to\infty} u(t,x) = \lim_{N\to\infty} \omega_{t,x}(m(\boldsymbol{\sigma}))$, we have the initial value problem

$$\begin{cases} \partial_t \bar{u}(t,x) + p\bar{u}(t,x)^{p-1}\partial_x \bar{u}(t,x) = 0, \\ \bar{u}(0,x) = \tanh(x), \end{cases} \tag{3.8}$$

where the initial profile is easily computed by straightforward calculations since it is a one-body problem. This system describes the propagation of nonlinear waves, and it can be solved by assuming a solution in implicit form $\bar{u}(t,x) = \tanh(x - v(t,x)t)$, where $v(t,x) = p\bar{u}(t,x)^{p-1}$ is the effective velocity. Recalling that the thermodynamics of the original p-spin model associated with the Hamilton function (3.1) is recovered by setting $t = -\beta$ and $x = 0$, we directly obtain

$$\bar{m} = \tanh(\beta p \bar{m}^{p-1}), \tag{3.9}$$

where $\bar{m} = \bar{u}(-\beta, 0)$. This is precisely the self-consistency equation for the global magnetization for the p-spin ferromagnetic model.[36] The phase transition of the system is expected to take place where the gradient of the solution explodes, which, on the Burgers side, corresponds to the development of a shock wave at $x = 0$. Since the temporal coordinate t is directly related to the thermal noise at which the phase transition occurs, with standard PDE methods, we can analytically determine the critical temperature according to the simple system

$$\begin{cases} \bar{\xi} = \dfrac{F(\bar{\xi})}{F'(\bar{\xi})}, \\ T_c = F'(\bar{\xi}), \end{cases}$$

where $F(\xi) = p\tanh(\xi)^{p-1}$. This prediction is in perfect agreement with the numerical solutions of the self-consistency Eq. (3.9).

IV. DERRIDA MODEL: HOW TO DEAL WITH COMPLEX SYSTEMS

In this section, we adapt the previous methodologies to treat complex systems with p-spin interactions. The paradigmatic case is given by the p-spin SK model, also referred to as Derrida model, defined below.

Definition 1. Let $\boldsymbol{\sigma}$ be the generic point in the configuration space $\Sigma_N \equiv \{-1, +1\}^N$ of the system. Let $\boldsymbol{J}$ be a p-rank random tensor with i.i.d. entries $J_{i_1\dots i_p} \sim \mathcal{N}(0,1)$. The Hamilton function of the p-spin Derrida model is defined as

$$H_{N,p,\boldsymbol{J}}(\boldsymbol{\sigma}) = -\sqrt{\frac{p!}{2N^{p-1}}} \sum_{1\le i_1<\dots<i_p\le N}^{N} J_{i_1\dots i_p}\sigma_{i_1}\dots\sigma_{i_p}. \tag{4.1}$$

Remark 1. For $p = 2$, we recover the Sherrington–Kirkpatrick model.[4]

Remark 2. In the usual definition of the p-spin SK model, the sum is performed with the constraint $1 \le i_1 < i_2 < \dots i_p \le N$ like in (4.1). Beyond that formulation, it is possible to consider an alternative one, where summation is realized independently over all the indices, the difference between the two expressions being vanishing in the thermodynamic limit, that is,

$$\sum_{1\le i_1<\dots<i_p\le N} (\cdot) = \frac{1}{p!}\sum_{i_1,\dots,i_p=1}^{N} (\cdot) + \text{ contributions vanishing as } N \to \infty. \tag{4.2}$$

Since we are interested in the thermodynamic limit, we will often use the equality

$$\sum_{1\le i_1<\dots<i_p\le N} (\cdot) = \frac{1}{p!}\sum_{i_1,\dots,i_p=1}^{N} (\cdot), \tag{4.3}$$

holding in the $N \to \infty$ limit.

Definition 2. Given $(t, x) \in \mathbb{R}^2$ and given a family $\{J_i\}_{i=1}^N$ of i.i.d. $\mathcal{N}(0,1)$-distributed random variables, Guerra's interpolating partition function for the p-spin SK model is

$$Z_{N,p,\boldsymbol{J}}(t,x) = \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp(-H_{N,p,\boldsymbol{J}}(t,x)), \tag{4.4}$$

$$H_{N,p,\boldsymbol{J}}(t,x) = -\sqrt{\frac{tp!}{2N^{p-1}}} \sum_{1\le i_1<\dots<i_p\le N}^{N} J_{i_1\dots i_p}\sigma_{i_1}\dots\sigma_{i_p} - \sqrt{x}\sum_{i=1}^{N} J_i\sigma_i. \tag{4.5}$$

The Boltzmann factor associated with this partition function is denoted with $B_{N,p,\boldsymbol{J}}(t,x)$.

As stated in Sec. I, when dealing with spin-glasses, we need to enlarge our analysis to the s-replicated version of the configuration space. To this aim, we use the following.

Definition 3. *Let $\Sigma_N^{(s)} = (\Sigma_N)^{\otimes s} \equiv \{-1,+1\}^{sN}$ be the s-replicated configuration space. We denote with $\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}^1,\ldots,\boldsymbol{\sigma}^s) \in \Sigma_N^{(s)}$ the global configuration of the replicated system. The space $\Sigma_N^{(s)}$ is naturally endowed with the s-replicated Boltzmann–Gibbs measure associated to the partition function*

$$Z_{N,p,\boldsymbol{J}}^{(s)}(t,x) = \sum_{\underline{\boldsymbol{\sigma}}\in\Sigma_N^{(s)}} \exp\left(\sqrt{\frac{tp!}{2N^{p-1}}}\sum_{a=1}^{s}\sum_{1\le i_1<\cdots<i_p\le N}^{N} J_{i_1\ldots i_p}\sigma_{i_1}^{(a)}\ldots\sigma_{i_p}^{(a)} + \sqrt{x}\sum_{a=1}^{s}\sum_{i=1}^{N} J_i\sigma_i^{(a)}\right). \tag{4.6}$$

We will denote with $B_{N,p,\boldsymbol{J}}^{(s)}(t,x)$ the Boltzmann factor appearing in the s-replicated partition function. Given an observable $O:\Sigma_N^{(s)}\to\mathbb{R}$ on the replicated space, the Boltzmann average with respect to the s-replicated partition function is

$$\Omega_{t,x}^{(s)}(O) = \frac{\sum_{\underline{\boldsymbol{\sigma}}} O(\underline{\boldsymbol{\sigma}}) B_{N,p,\boldsymbol{J}}^{(s)}(t,x)}{\sum_{\underline{\boldsymbol{\sigma}}} B_{N,p,\boldsymbol{J}}^{(s)}(t,x)}. \tag{4.7}$$

Remark 3. *The thermodynamics of the original model is recovered with $t=\beta^2$ and $x=0$.*

Remark 4. *Since the replicas are independent, $Z_{N,p,\boldsymbol{J}}^{(s)}(t,x) \equiv (Z_{N,p,\boldsymbol{J}}(t,x))^s$.*

In the following, in order to lighten the notation, the replica index s of the Boltzmann average $\Omega_{t,x}^{(s)}$ can be dropped, since it is understood directly from the function to be averaged.

Definition 4. *Given an observable $O:\Sigma_N^{(s)}\to\mathbb{R}$ on the replicated space, the quenched average is defined as*

$$\langle O\rangle_{t,x} = \mathbb{E}_{\boldsymbol{J}}\Omega_{t,x}(O). \tag{4.8}$$

Remark 5. *In the last definition, the average $\mathbb{E}_{\boldsymbol{J}}$ is again the expectation value performed over all the quenched disorder, thus including the auxiliary random variables in the interpolating setup.*

Definition 5. *The order parameter for the p-spin SK model is the replica overlap*

$$q_{ab} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^{(a)}\sigma_i^{(b)}, \tag{4.9}$$

where $\boldsymbol{\sigma}^{(a)}$ and $\boldsymbol{\sigma}^{(b)}$ are two generic configurations of different replicas of the system labeled, respectively, a and b.

We can now focus on the PDE approach to the statistical mechanics of the p-spin SK model. The final goal is to prove the following theorem.

Theorem 1. *The expectation value of the order parameter for the p-spin Derrida model under the RS ansatz is given by the function $\bar{q}(\beta) = u(\beta^2,0)$, where $u(t,x)$ is the solution of the inviscid limit of the p-th element Burgers hierarchy with initial profile (4.32), i.e.,*

$$\begin{cases} \partial_t u(t,x) - \frac{1}{2}\partial_x u^p(t,x) = 0, \\ u(0,x) = \mathbb{E}_J \tanh^2(\sqrt{x}J). \end{cases} \tag{4.10}$$

To this aim, it is necessary to first prove some preliminary properties, as detailed below.

Definition 6. *For all $p\ge 2$, Guerra's action functional is defined as*

$$S_{N,p}(t,x) = 2A_{N,p}(t,x) - x - \frac{t}{2}. \tag{4.11}$$

Lemma 1. *The spacetime derivatives of Guerra's action functional read as follows:*

$$\partial_t S_{N,p}(t,x) = -\frac{1}{2}\langle q_{12}^p\rangle_{t,x} + R_N(t,x), \tag{4.12}$$

$$\partial_x S_{N,p}(t,x) = -\langle q_{12}\rangle_{t,x}, \tag{4.13}$$

where $R_N(t,x)$ takes into account the contributions coming from (4.2) and vanishing in the $N \to \infty$ limit.

The proof of this lemma can be found in Appendix A.

Lemma 2. Given an observable $O : \Sigma_N^{(s)} \to \mathbb{R}$ on the replicated space, the following streaming equation holds:

$$\partial_x \langle O(\underline{\boldsymbol{\sigma}})\rangle_{t,x} = \frac{N}{2}\sum_{a,b=1}^{s}\langle O(\underline{\boldsymbol{\sigma}})q_{ab}\rangle_{t,x} - sN\sum_{a=1}^{s}\langle O(\underline{\boldsymbol{\sigma}})q_{a,s+1}\rangle_{t,x} - \frac{s}{2}N\langle O(\underline{\boldsymbol{\sigma}})\rangle_{t,x} + \frac{s(s+1)}{2}\langle O(\underline{\boldsymbol{\sigma}})q_{s+1,s+2}\rangle_{t,x}. \tag{4.14}$$

Proof. The proof is long and rather cumbersome, so we will just give a sketch. First of all, we recall that

$$\langle O(\underline{\boldsymbol{\sigma}})\rangle_{t,x} = \mathbb{E}_{\boldsymbol{J}}\frac{1}{Z_{N,p,\boldsymbol{J}}(t,x)^s}\sum_{\underline{\boldsymbol{\sigma}}\in\Sigma_N^{(s)}} O(\underline{\boldsymbol{\sigma}})B^{(s)}_{N,p,\boldsymbol{J}}(t,x). \tag{4.15}$$

When taking the x-derivative of this quantity, we will get two contributions: the first one follows from the derivative of $B^{(s)}_{N,p,\boldsymbol{J}}(t,x)$ and the second one follows from the derivative of $1/Z^s_{N,p,\boldsymbol{J}}$ (which results in adding a new replica). In quantitative terms,

$$\partial_x\langle O(\underline{\boldsymbol{\sigma}})\rangle_{t,x} = \frac{1}{2\sqrt{x}}\mathbb{E}_{\boldsymbol{J}}\Bigg(\sum_{a=1}^{s}\sum_{i=1}^{N} J_i\frac{1}{Z_{N,p,\boldsymbol{J}}(t,x)^s}\sum_{\boldsymbol{\sigma}^{(1)}}\dots\sum_{\boldsymbol{\sigma}^{(s)}} O(\underline{\boldsymbol{\sigma}})\sigma_i^{(a)}B^{(s)}_{N,p,\boldsymbol{J}}(t,x) - s\sum_{i=1}^{N} J_i\frac{1}{Z_{N,p,\boldsymbol{J}}(t,x)^{s+1}}\sum_{\boldsymbol{\sigma}^{(1)}}\dots\sum_{\boldsymbol{\sigma}^{(s+1)}} O(\underline{\boldsymbol{\sigma}})\sigma_i^{(s+1)}B^{(s+1)}_{N,p,\boldsymbol{J}}(t,x)\Bigg). \tag{4.16}$$

The presence of J_i in both terms of the right-hand side can be carried out by applying the Wick–Isserlis theorem. Then, each J_i-derivative results in two different contributions that can be recast as the difference of two s-replicated Boltzmann averages (in fact, notice that the action of the derivative on the denominators involving the partition functions results in the appearance of a replicated Boltzmann factor). Furthermore, the explicit x-dependence of the derivative precisely cancels out (since the J_i-derivative will produce factors proportional to $\sqrt{x}$). After all these passages, and remembering $\langle\cdot\rangle_{t,x} = \mathbb{E}_{\boldsymbol{J}}\Omega_{t,x}(\cdot)$, we get

$$\partial_x\langle O(\underline{\boldsymbol{\sigma}})\rangle_{t,x} = \frac{1}{2}\sum_{a,b=1}^{s}\sum_{i=1}^{N}\langle O(\underline{\boldsymbol{\sigma}})\sigma_i^{(a)}\sigma_i^{(b)}\rangle_{t,x} - \frac{s}{2}\sum_{a=1}^{s}\sum_{i=1}^{N}\langle O(\underline{\boldsymbol{\sigma}})\sigma_i^{(a)}\sigma_i^{(s+1)}\rangle_{t,x} - \frac{s}{2}\sum_{s=1}^{s+1}\sum_{i=1}^{N}\langle O(\underline{\boldsymbol{\sigma}})\sigma_i^{(a)}\sigma_i^{(s+1)}\rangle_{t,x} + \frac{s(s+1)}{2}\sum_{i=1}^{N}\langle O(\underline{\boldsymbol{\sigma}})\sigma_i^{(s+1)}\sigma_i^{(s+2)}\rangle_{t,x}. \tag{4.17}$$

Recalling Definition 5 and after some rearrangements of the quantities, we get the thesis. □

In order to proceed, we have now to make some physical assumptions on the model. As standard in spin-glass theory, the simplest requirement is the RS in the thermodynamic limit. In fact, as we are going to show, this makes the PDE approach feasible due to the fact that we can express nontrivial expectation values of functions of the replicas in a very simple form.

Proposition 1. For the interpolated Derrida model (4.4) and (4.5), the following equality holds:

$$\langle q_{12}^p\rangle_{t,x} = \left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,x}\right)^{p-1}\langle q_{12}\rangle_{t,x} + Q_N^{(p-1)}(t,x), \tag{4.18}$$

where $Q_N^{(p-1)}(t,x)$ vanishes in the $N \to \infty$ limit and under the RS assumption.

Proof. Let us consider the x-derivative of $\langle q_{12}\rangle$ and try to rearrange the first contribution as follows:

$$\langle q_{12}^l q_{23}\rangle_{t,x} = \langle q_{12}^l\Delta(q_{23})\rangle_{t,x} + \langle q_{23}\rangle_{t,x}\langle q_{12}^l\rangle_{t,x}, \tag{4.19}$$

where $\Delta(q_{ab}) = q_{ab} - \langle q_{ab}\rangle_{t,x}$ $\forall a,b$ is the fluctuation of the overlap with respect to its thermodynamic value. Furthermore,

$$\begin{aligned}\langle q_{12}^l q_{23}\rangle_{t,x} &= \langle \Delta(q_{12}^l)\Delta(q_{23})\rangle_{t,x} + \langle q_{12}^l\rangle_{t,x}\langle \Delta(q_{23})\rangle_{t,x} + \langle q_{23}\rangle_{t,x}\langle q_{12}^l\rangle_{t,x}\\ &= \langle q_{12}\rangle_{t,x}\langle q_{12}^l\rangle_{t,x} + R_N^{(1,l)}(t,x),\end{aligned} \tag{4.20}$$

where $R_N^{(1,l)}(t,x)$ represents the terms involving the fluctuation functions of the overlap. In the last equality, we also used the fact that $\langle q_{23}\rangle_{t,x}$ $= \langle q_{12}\rangle_{t,x}$ since the average is independent of the labeling of replicas.

We now notice that, by using $O(\underline{\sigma}) = q_{12}^l$ in Lemma 2, we reach the following equality for each $l \in \mathbb{N}$:

$$\partial_x\langle q_{12}^l\rangle_{t,x} = N\Big(\langle q_{12}^{l+1}\rangle_{t,x} - 4\langle q_{12}^l q_{23}\rangle_{t,x} + 3\langle q_{12}^l q_{34}\rangle_{t,x}\Big), \tag{4.21}$$

and for the last contribution in the right-hand side, we can apply an expansion analogous to (4.20),

$$\langle q_{12}^l q_{34}\rangle_{t,x} = \langle q_{12}\rangle_{t,x}\langle q_{12}^l\rangle_{t,x} + R_N^{(2,l)}(t,x). \tag{4.22}$$

Thus, combining (4.20)–(4.22), we finally get

$$\langle q_{12}^{l+1}\rangle_{t,x} = \frac{1}{N}\partial_x\langle q_{12}^l\rangle_{t,x} + \langle q_{12}\rangle_{t,x}\langle q_{12}^l\rangle_{t,x} + R_N^{(l)}(t,x), \tag{4.23}$$

where $R_N^{(l)}(t,x) = 4R_N^{(1,l)}(t,x) - 3R_N^{(2,l)}(t,x)$. We can then express higher moments of the overlap in terms of lower ones,

$$\langle q_{12}^{l+1}\rangle_{t,x} = \left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,x}\right)\langle q_{12}^l\rangle_{t,x} + R_N^{(l)}(t,x). \tag{4.24}$$

Iterating this procedure from $l = 1$ up to $l = p-1$, we obtain

$$\langle q_{12}^p\rangle_{t,x} = \left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,x}\right)^{p-1}\langle q_{12}\rangle_{t,x} + Q_N^{(p-1)}(t,x), \tag{4.25}$$

where $Q_N^{(p-1)}(t,x)$ collects all the terms involving $R_N^{(l)}(t,x)$ and, thus, vanishes in the $N\to\infty$ limit. □

At this point, we have at our disposal all the ingredients needed for making explicit our approach. Using (4.25) in (4.12), we get

$$\partial_t S_{N,p}(t,x) = -\frac{1}{2}\left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,x}\right)^{p-1}\langle q_{12}\rangle_{t,x} + R_N(t,x) - \frac{1}{2}Q_N^{(p-1)}(t,x). \tag{4.26}$$

Deriving (4.26) with respect to the spatial coordinate x, we have

$$\partial_t\partial_x S_{N,p}(t,x) = -\frac{1}{2}\partial_x\left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,x}\right)^{p-1}\langle q_{12}\rangle_{t,x} + V_N(t,x),$$

where $V_N(t,x) := -\partial_x(R_N(t,x) - \frac{1}{2}Q_N^{(p-1)}(t,x))$ is vanishing in the $N\to\infty$ limit. Then, recalling Eq. (4.13), we can write the following equation:

$$\partial_t\langle q_{12}\rangle_{t,x} - \frac{1}{2}\partial_x\left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,x}\right)^{p-1}\langle q_{12}\rangle_{t,x} = V_N(t,x). \tag{4.27}$$

On the lhs, we recognize a Burgers hierarchy structure, while on the rhs, we have a source term (which further vanishes in the thermodynamic limit).

We are finally ready to prove Theorem 1, as reported hereafter.

Proof 1. By taking the limit of Eq. (4.27) for $N\to\infty$ and recalling that $V_N(t,x)\to 0$ for $N\to\infty$, we get

$$\partial_t u(t,x) - \frac{1}{2}\partial_x u^p(t,x) = 0, \tag{4.28}$$

where $u(t,x) := \lim_{N\to\infty} \langle q_{12} \rangle_{t,x}$. The initial profile of the Cauchy problem associated with the PDE (4.28) is easily determined, since for $t = 0$, the partition function reduces to a one-body problem. Thus, we have to compute $u(0,x) = \lim_{N\to\infty} \langle q_{12}\rangle_{0,x}$. To achieve this aim, we start from the partition function evaluated at $t = 0$, which is

$$Z_{N,p,\boldsymbol{J}}(0,x) = \sum_{\boldsymbol{\sigma}} \exp\left(\sqrt{x}\sum_{i=1}^{N} J_i\sigma_i\right) = \prod_{i=1}^{N} 2\cosh(\sqrt{x}J_i). \tag{4.29}$$

Taking the logarithm and averaging over the quenched disorder $\boldsymbol{J}$, we have the intensive pressure,

$$A_{N,p}(0,x) = \frac{1}{N}\mathbb{E}_{\boldsymbol{J}} \log \prod_{i=1}^{N} 2\cosh(\sqrt{x}J_i) = \log 2 + \frac{2}{N}\sum_{i=1}^{N}\mathbb{E}_{\boldsymbol{J}} \log\cosh(\sqrt{x}J_i). \tag{4.30}$$

Recalling that Guerra's action is defined as in (4.11) and that the J_i are i.i.d., so the sum of quenched averages of functions of J_i is N times the average with respect to a *single* quenched variable $J \sim \mathcal{N}(0,1)$, we get

$$S_N(0,x) = 2\log 2 + 2\mathbb{E}_J \log\cosh(\sqrt{x}J) - x. \tag{4.31}$$

Finally, taking the derivative with respect to the spatial coordinate, we have the initial profile for the overlap expectation value, which reads

$$u(0,x) = \lim_{N\to\infty} \langle q_{12}\rangle_{0,x} = \mathbb{E}_J \tanh^2(\sqrt{x}J). \tag{4.32}$$

Here, we again used the Wick–Isserlis theorem for normally distributed random variables. Putting together (4.28) and (4.10), we get the thesis. □

Corollary 1. The implicit solution of the inviscid Burgers hierarchy (4.10) is the self-consistency equation for the order parameter $\bar{q}(\beta)$ for the p-spin model under the RS ansatz.

Proof. Let us rewrite the differential Eq. (4.10) as

$$\partial_t u - \frac{p}{2}u^{p-1}\partial_x u = 0. \tag{4.33}$$

This is a nonlinear wave equation and, as well-known, it admits a solution of the form $u(t,x) = u_0(x - v(t,x)t)$, where u_0 is the initial profile and $v(t,x)$ is the effective velocity. For the case under consideration, we have $v(t,x) = -\frac{p}{2}u^{p-1}(t,x)$; thus,

$$u(t,x) = \mathbb{E}_J \tanh^2\left(\sqrt{x + t\frac{p}{2}u(t,x)^{p-1}}J\right). \tag{4.34}$$

Recalling that $\bar{q}(\beta) = u(\beta^2, 0)$, we finally have

$$\bar{q} = \mathbb{E}_J \tanh^2\left(\beta\sqrt{\frac{p}{2}\bar{q}^{p-1}}J\right), \tag{4.35}$$

which is precisely the self-consistency equation for the p-spin glass model, as reported also in Ref. 27. □

Corollary 2. The (ergodicity breaking) phase transition of the p-spin model coincides with the gradient catastrophe of the Cauchy problem (4.10), and the critical temperature is determined by the system parameters,

$$\begin{cases} T_c = \sqrt{-F'(\bar{\xi})}, \\ \bar{\xi} = \dfrac{F(\bar{\xi})}{F'(\bar{\xi})}, \end{cases} \tag{4.36}$$

where $F(\xi) = -\frac{p}{2}\mathbb{E}_J \tanh^2(\sqrt{x}J)$.

Proof. The determination of the critical temperature can be achieved with the usual analysis of intersecting characteristics of the Cauchy problem (4.10), and follows the same lines of Ref. 36. □

As a comparison, in Fig. 1, we reported the solutions of the self-consistency Eq. (4.35) for $p = 2, \ldots, 8$ (solid curves) and the critical temperatures as predicted by the system (4.36) (dashed lines).

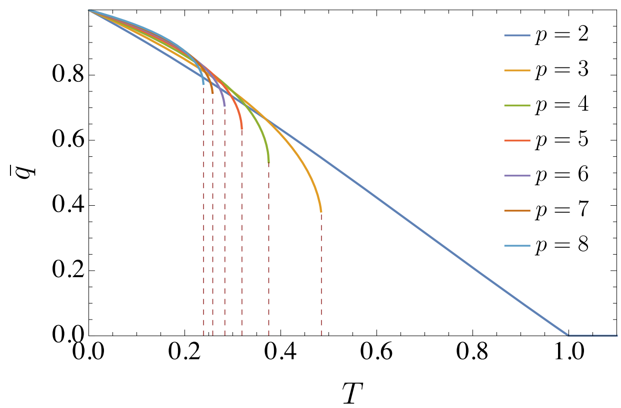

FIG. 1. Solutions of the self-consistency Eq. (4.35) for $p = 2, \ldots, 8$ (solid curves) and the critical temperatures as predicted by the system (4.36) (dashed lines).

V. APPLICATION TO DENSE ASSOCIATIVE MEMORIES

Going beyond the pure spin-glass case, in this section, we will approach the DAMs that constitute a generalization of the Hopfield model where neurons are embedded on hyper-graphs of order p, in such a way that they interact in groups of p units and the Hopfield reference is recovered for $p = 2$. These systems are the main focus of this work.

Definition 7. Let $\boldsymbol{\sigma}$ be the generic point in the configuration space $\Sigma_N \equiv \{-1, +1\}^N$ of the system. Given K random patterns $\{\boldsymbol{\xi}^\mu\}_{\mu=1}^K$ each made of N i.i.d. binary entries drawn with equal probability $P(\xi_i^\mu = -1) = P(\xi_i^\mu = 1) = \frac{1}{2}$ $\forall i = 1, \ldots, N$, the Hamiltonian of the p-th order DAM is

$$H_{N,p,\xi,K}(\boldsymbol{\sigma}) := -\frac{1}{N^{p-1}} \sum_{\mu=1}^{K} \sum_{i_1,\ldots,i_p}^{N} \xi_{i_1}^\mu \ldots \xi_{i_p}^\mu \sigma_{i_1} \ldots \sigma_{i_p}. \tag{5.1}$$

Remark 6. The normalization factor $\frac{1}{N^{p-1}}$ ensures the linear extensivity of the Hamiltonian, in the volume of the network N, i.e., $\lim_{N\to\infty} \left| \frac{H_{N,p,\xi,K}}{N} \right| \in (0, +\infty)$.

As anticipated in Sec. I, DAMs have been proved to exhibit high computational skills.[21–23] Among these, we recall that they are able to store a number of patterns K scaling as $K \sim N^{p-1}$[22,23] and that, if supplied with a relatively small number of patterns, they can retrieve them also in the presence of additional and extensive sources of noise.[21,49] In general, the network performance can be split into two regimes: (1) the low-load one, where $\lim_{N\to\infty} \frac{K}{N^{p-1}} = 0$, and (2) the high-storage one, where $\lim_{N\to\infty} \frac{K}{N^{p-1}} > 0$; in the following, we will address both of them.

A. Low storage

Let us begin the analysis of the network in a low-load regime, setting the number of stored patterns as finite. Again, the goal is to use interpolation techniques and derive PDEs able to describe the thermodynamics of the system. To do this, let us start by defining the below quantity.

Definition 8. The order parameters used to describe the macroscopic behavior of the model are the so-called Mattis magnetizations, defined as

$$m_\mu(\boldsymbol{\sigma}) := \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \sigma_i \quad \forall \mu = 1, \ldots, K, \tag{5.2}$$

measuring the overlap between the network configuration and the stored patterns.

Remark 7. The Hamilton function (5.1) in terms of the Mattis magnetizations is

$$H_{N,p,\xi,K}(\boldsymbol{\sigma}) = -N \sum_{\mu=1}^{K} (m_\mu(\boldsymbol{\sigma}))^p.$$

Next, we define the basic objects of our investigations within the interpolating framework.

Definition 9. Given $(t, \boldsymbol{x}) \in \mathbb{R}^{K+1}$, *the spacetime Guerra's interpolating partition function for the DAM model (in the low-load regime) reads as*

$$Z_{N,p,\xi,K}(t, \boldsymbol{x}) = \sum_{\boldsymbol{\sigma} \in \Sigma_N} \exp\big(-H_{N,p,\xi,K}(t, \boldsymbol{x})\big), \tag{5.3}$$

$$H_{N,p,\xi,K}(t, \boldsymbol{x}) = tN \sum_{\mu=1}^{K} m_\mu(\boldsymbol{\sigma})^p - N \sum_{\mu=1}^{K} x_\mu m_\mu(\boldsymbol{\sigma}). \tag{5.4}$$

Remark 8. The spacetime Guerra's interpolating partition function recovers the one related to the DAM by setting $t = -\beta$ *and* $\boldsymbol{x} = \boldsymbol{0}$.

We recall that, for the CW model, the Guerra mechanical analogy consists in interpreting the statistical pressure as the Burgers hierarchy describing the motion of viscid nonlinear waves in 1 + 1-dimensional space. In the case of the DAMs, we have K Mattis magnetizations, and the dual mechanical system describes nonlinear waves traveling in a K + 1-dimensional space.

Definition 10. For each configuration $\boldsymbol{\sigma} \in \Sigma_N$ *of the system, the Boltzmann factor corresponding to the partition function (5.3) is*

$$B_{N,p,\xi,K}(t, \boldsymbol{x}) = \exp\left(-tN \sum_{\mu=1}^{K} m_\mu(\boldsymbol{\sigma})^p + N \sum_{\mu=1}^{K} x_\mu m_\mu(\boldsymbol{\sigma})\right). \tag{5.5}$$

Lemma 3. The first-order spacetime derivatives of the Guerra intensive pressure associated to the partition function (5.3) read as

$$\partial_t A_{N,p,\xi,K}(t, \boldsymbol{x}) = -\sum_{\mu=1}^{K} \omega_{t,\boldsymbol{x}}(m_\mu(\boldsymbol{\sigma})^p), \tag{5.6}$$

$$\partial_\mu A_{N,p,\xi,K}(t, \boldsymbol{x}) = \omega_{t,\boldsymbol{x}}(m_\mu(\boldsymbol{\sigma})), \tag{5.7}$$

where $\partial_\mu := \partial_{x^\mu}$.

Proof. Recalling the definition of the intensive pressure $A_{N,p,\xi,K}(t, \boldsymbol{x}) = \frac{1}{N} \log \sum_{\boldsymbol{\sigma} \in \Sigma_N} B_{N,p,\xi,K}(t, \boldsymbol{x})$, along with Eq. (5.5), the proof follows straightforward computations. The temporal derivative reads

$$\partial_t A_{N,p,\xi,K}(t, \boldsymbol{x}) = \frac{1}{N} Z^{-1}_{N,p,\xi,K}(t, \boldsymbol{x}) \sum_{\boldsymbol{\sigma} \in \Sigma_N} \left(-N \sum_{\mu=1}^{K} m_\mu(\boldsymbol{\sigma})^p\right) B_{N,p,\xi,K}(t, \boldsymbol{x}) = -\sum_{\mu=1}^{K} \omega_{t,\boldsymbol{x}}(m_\mu(\boldsymbol{\sigma})^p),$$

while the spatial derivative reads

$$\partial_\mu A_{N,p,\xi,K}(t, \boldsymbol{x}) = \frac{1}{N} Z^{-1}_{N,p,\xi,K}(t, \boldsymbol{x}) \sum_{\boldsymbol{\sigma} \in \Sigma_N} (N m_\mu(\boldsymbol{\sigma})) B_{N,p,\xi,K}(t, \boldsymbol{x}) = \omega_{t,\boldsymbol{x}}(m_\mu(\boldsymbol{\sigma})).$$

□

Proposition 2. The higher (non-centered) momenta of the Mattis magnetizations are realized as

$$\omega_{t,\boldsymbol{x}}(m_\mu^{s+1}) = \left(\frac{1}{N} \partial_\mu + \omega_{t,\boldsymbol{x}}(m_\mu)\right) \omega_{t,\boldsymbol{x}}(m_\mu^s), \tag{5.8}$$

for each integer $s \geq 1$.

Proof. We start by computing the spatial derivative of the Mattis magnetizations expectation value,

$$\begin{aligned}
\partial_\nu \omega_{t,\boldsymbol{x}}(m_\mu^s) &= \partial_\nu \left(Z^{-1}_{N,p,\xi,K}(t, \boldsymbol{x}) \sum_{\boldsymbol{\sigma} \in \Sigma_N} m_\mu(\boldsymbol{\sigma})^s B_{N,p,\xi,K}(t, \boldsymbol{x})\right) \\
&= N Z^{-1}_{N,p,\xi,K}(t, \boldsymbol{x}) \sum_{\boldsymbol{\sigma} \in \Sigma_N} m_\mu(\boldsymbol{\sigma})^s m_\nu(\boldsymbol{\sigma}) B_{N,p,\xi,K}(t, \boldsymbol{x}) \\
&\quad - N Z^{-1}_{N,p,\xi,K}(t, \boldsymbol{x}) \sum_{\boldsymbol{\sigma} \in \Sigma_N} m_\mu(\boldsymbol{\sigma})^s B_{N,p,\xi,K}(t, \boldsymbol{x}) \cdot Z^{-1}_{N,p,\xi,K}(t, \boldsymbol{x}) \sum_{\boldsymbol{\sigma}' \in \Sigma_N} m_\nu(\boldsymbol{\sigma}') B_{N,p,\xi,K}(t, \boldsymbol{x}) \\
&= N \omega_{t,\boldsymbol{x}}(m_\mu^s m_\nu) - N \omega_{t,\boldsymbol{x}}(m_\mu^s) \omega_{t,\boldsymbol{x}}(m_\nu).
\end{aligned}$$

In particular, for $\nu = \mu$, we have

$$\partial_\mu \omega_{t,\boldsymbol{x}}(m_\mu^s) = N[\omega_{t,\boldsymbol{x}}(m_\mu^{s+1}) - \omega_{t,\boldsymbol{x}}(m_\mu^s)\omega_{t,\boldsymbol{x}}(m_\mu)].$$

Expressing the higher order moment in terms of the other quantities, we reach the thesis. □

By calling $u_\mu^{(s)}(t,\boldsymbol{x}) := \omega_{t,\boldsymbol{x}}(m_\mu(\boldsymbol{\sigma})^s)$, we can express all $u_\mu^{(s)}(t,\boldsymbol{x})$ in terms of $u_\mu^{(1)}(t,\boldsymbol{x}) := u_\mu(t,\boldsymbol{x})$ for each $s > 1$. Indeed,

$$u_\mu^{(s+1)}(t,\boldsymbol{x}) = \left(\frac{1}{N}\partial_\mu + u_\mu(t,\boldsymbol{x})\right)u_\mu^{(s)}(t,\boldsymbol{x}) = \left(\frac{1}{N}\partial_\mu + u_\mu(t,\boldsymbol{x})\right)^s u_\mu(t,\boldsymbol{x}). \tag{5.9}$$

To simplify the notation, we define the operator $D_\mu := \frac{1}{N}\partial_\mu + u_\mu(t,\boldsymbol{x})$.

Theorem 2. *The expectation value of the Mattis magnetizations of the interpolated DAM model (5.3) and (5.4) satisfies the nonlinear evolutive equations given by*

$$\partial_t u_\mu(t,\boldsymbol{x}) = -\sum_{\nu=1}^K \partial_\mu D_\nu^{p-1} u_\nu(t,\boldsymbol{x}). \tag{5.10}$$

Proof. First, we put $s = p-1$ in (5.9) so that

$$u_\nu^{(p)}(t,\boldsymbol{x}) = D_\nu^{p-1} u_\nu(t,\boldsymbol{x}).$$

Now, recall that $u_\nu^{(p)}(t,\boldsymbol{x}) = \omega_{t,\boldsymbol{x}}(m_\nu(\boldsymbol{\sigma})^p)$ and $\partial_t A_{N,p,\xi,K}(t,\boldsymbol{x}) = -\sum_{\mu=1}^K \omega_{t,\boldsymbol{x}}(m_\mu(\boldsymbol{\sigma})^p)$; thus,

$$\partial_t A_{N,p,\xi,K}(t,\boldsymbol{x}) = -\sum_{\nu=1}^K u_\nu^{(p)}(t,\boldsymbol{x}) = -\sum_{\nu=1}^K D_\nu^{p-1} u_\nu(t,\boldsymbol{x}).$$

Taking the derivative ∂_μ, commuting ∂_t and ∂_μ and recalling that $\partial_\mu A_{N,p,\xi,K}(t,\boldsymbol{x}) = \omega_{t,\boldsymbol{x}}(m_\mu(\boldsymbol{\sigma})) = u_\mu(t,\boldsymbol{x})$, we directly reach the thesis. □

Before proceeding, it is useful to recall that the evolutive Eq. (5.10) can be linearized by means of the Cole–Hopf transform. In fact, performing the multidimensional Cole–Hopf transform $u_\mu(t,\boldsymbol{x}) = \frac{1}{N}\partial_\mu(\log\Psi)$, we have

$$\frac{1}{N}\partial_t \frac{\partial_\mu\Psi}{\Psi} = -\frac{1}{N^p}\sum_{\nu=1}^K \partial_\mu\left(\partial_\mu + \frac{\partial_\mu\Psi}{\Psi}\right)^{p-1}\frac{\partial_\mu\Psi}{\Psi},$$

and by using the basic identities

$$\left(\partial_\mu + \frac{\partial_\mu\Psi}{\Psi}\right)^s \frac{\partial_\mu\Psi}{\Psi} = \frac{\partial_\mu^{s+1}\Psi}{\Psi},$$

$$\partial_t\frac{\partial_\mu\Psi}{\Psi} = \partial_\mu\frac{\partial_t\Psi}{\Psi},$$

we have

$$\partial_\mu\left(\frac{\partial_t\Psi}{\Psi} + \sum_{\nu=1}^K \frac{\partial_\mu^p\Psi}{\Psi}\right) = 0.$$

Setting the argument of the spatial derivative to zero and assuming $\Psi \neq 0$, we have

$$\partial_t\Psi + \frac{1}{N^{p-1}}\sum_{\nu=1}^K \partial_\nu^p\Psi = 0.$$

Remark 9. In the proof, the function Ψ is nothing but Guerra's interpolating partition function, as can be understood by comparing the definitions $u_\mu(t,\boldsymbol{x}) = \frac{1}{N}\partial_\mu \log\Psi(t,\boldsymbol{x})$ and $u_\mu(t,\boldsymbol{x}) = \partial_\mu A_{N,p,\xi,K}(t,\boldsymbol{x})$. Indeed, by computing the derivatives of the partition function, we easily get

$$\partial_t Z_{N,p,\xi,K}(t,\boldsymbol{x}) = -N\sum_{\mu=1}^K \sum_{\sigma\in\Sigma_N} m_\mu^p(\boldsymbol{\sigma}) B_{N,p,\xi,K}(t,\boldsymbol{x}),$$

$$\partial_\mu^p Z_{N,p,\xi,K}(t,\boldsymbol{x}) = N^p \sum_{\sigma\in\Sigma_N} m_\mu^p(\boldsymbol{\sigma}) B_{N,p,\xi,K}(t,\boldsymbol{x}).$$

A direct comparison shows that Guerra's interpolating partition function satisfies the same differential equation of the Ψ potential.

Remark 10. The case $K = 1$ corresponds to the p-spin CW model treated in Ref. 36. Indeed, the partition function of the system can be handled as

$$Z_{N,p,\xi,K=1}(t,\boldsymbol{x}) = \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp\left(-tN\left(\frac{1}{N}\sum_i \xi_i^1\sigma_i\right)^p + Nx\left(\frac{1}{N}\sum_i \xi_i^1\sigma_i\right)\right)$$
$$= \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp\left(-tN\left(\frac{1}{N}\sum_i \sigma_i\right)^p + Nx\left(\frac{1}{N}\sum_i \sigma_i\right)\right) = Z_{N,p}^{(CW)}(t,\boldsymbol{x}),$$

where we used the invariance of the partition function under the transformation $\sigma_i \to \xi_i^1\sigma_i$. In this particular case, we recover the Burgers hierarchy with viscosity parameter $1/N$: calling $x_1 = x$ and $u_1(t,\boldsymbol{x}) = u(t,\boldsymbol{x})$, family (5.10) reduces to

$$\partial_t u + \partial_x\left(\frac{1}{N}\partial_x + u\right)^{p-1} u = 0.$$

Within this framework, we generate the multidimensional generalization of Burgers hierarchy, see Appendix B for further details and examples.

B. High storage

Here, we will study the p-spin DAMs in the high-load regime $\lim_{N\to\infty}\frac{K}{N^{p-1}} = \alpha > 0$ for even p, which, unlike the low-load regime just discussed, exhibits a complex behavior. Let us start by observing that, in this case, the partition function related to the Hamiltonian (5.1) can also be written in the following form (notice the little abuse of notation in the expression $Z_{N,p,\xi,\alpha}(\beta)$: in the subscript, α is meant as the ratio K/N^{p-1} by-passing the thermodynamic limit):

$$Z_{N,p,\xi,\alpha}(\beta) := \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp(-\beta H_{N,p,\xi,\alpha}(\boldsymbol{\sigma}))$$
$$= \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp\left[\frac{\beta}{p!N^{p-1}}\sum_{\mu=1}^{K}\sum_{i_1,\ldots,i_p}^{N} \xi_{i_1}^\mu \ldots \xi_{i_p}^\mu \sigma_{i_1}\ldots\sigma_{i_p}\right]. \tag{5.11}$$

Here, we used the index α rather than K in order to distinguish between the partition function of low- and high-storage regimes.

As standard in statistical mechanics of (complex) neural networks, we will assume that a single pattern is the candidate to be retrieved, say $\boldsymbol{\xi}^1$. Under this assumption, we can treat separately the Mattis magnetization $m := m_1$ corresponding to the recalled pattern from those associated with non-retrieved ones. Accordingly, in the partition function (5.11), the sum over μ is split into a signal contribution (corresponding to $\mu = 1$) and a noise contribution (corresponding to $\mu > 1$). For the latter, we apply a generalization of the quenched-noise universality property,[50–52] whence

$$\sum_{\mu\geq 2}^{K}\sum_{i_1,\ldots,i_p}^{N} \xi_{i_1}^\mu \ldots \xi_{i_p}^\mu \sigma_{i_1}\ldots\sigma_{i_p} \sim \mathcal{V}\sum_{\mu\geq 2}^{K}\left(\sum_{i_1<\cdots<i_{p/2}} J^\mu_{i_1,\ldots,i_{p/2}}\sigma_{i_1}\ldots\sigma_{i_{p/2}}\right)^2, \tag{5.12}$$

where $J^\mu_{i_1,\ldots,i_{p/2}} \sim_{iid} \mathcal{N}(0,1)$ are effective interaction strengths and the parameter $\mathcal{V}$ controls the scale of these effective interactions. In the expression above, the symbol ~ means that the two sides are Gaussian random variables in the thermodynamic limit, i.e., $N, K \to \infty$, with the same second moment (the first moment is unessential since it would contribute to the free energy as $\mathcal{O}(\log N/N)$). The quantity $\mathcal{V}$ can be computed as detailed in Appendix C, and it reads

$$\mathcal{V} = \frac{p}{2}!\sqrt{\frac{(2p-1)!! - ((p-1)!!)^2}{2}}. \tag{5.13}$$

For later convenience, we define $\mathcal{V} = \frac{p}{2}!\mathcal{V}_0$ (notice that, for $p = 2$, we have $\mathcal{V}_0 = 1$). With this definition, we can directly replace the noisy contribution in the partition function with the rhs of (5.12) and perform a Hubbard–Stratonovich transformation.

Definition 11. Given $(t,\boldsymbol{x}) \in \mathbb{R}^4$, the spacetime Guerra's interpolating partition function for the DAM model in the high-load regime reads as

$$Z_{N,p,\xi,\alpha}(t,\boldsymbol{x}) = \sum_{\boldsymbol{\sigma}\in\Sigma_N}\int\left(\prod_{\mu=1}^{K} d\tau_\mu \frac{e^{-\frac{\tau_\mu^2}{2}}}{\sqrt{2\pi}}\right)\exp(-H_{N,p,\xi,\alpha}(t,\boldsymbol{x})), \tag{5.14}$$

$$
\begin{aligned}
H_{N,p,\xi,\alpha}(t,\boldsymbol{x}) = &-\frac{tN}{2}m^p - \sqrt{\frac{t}{N^{p-1}}}\sqrt{\mathcal{V}}\sum_{\mu=2}^{K}\sum_{i_1<\cdots<i_{p/2}}^{N} J^{\mu}_{i_1,\ldots,i_{p/2}}\sigma_{i_1}\ldots\sigma_{i_{p/2}}\tau_\mu \\
&- \sqrt{x}\sum_{i=1}^{N}\eta_i\sigma_i - \sqrt{N^{1-p/2}y}\sum_{\mu=1}^{K}\theta_\mu\tau_\mu + \frac{N^{1-p/2}}{2}[\mathcal{V}_0(t-t_0)+y]\sum_{\mu=1}^{K}\tau_\mu^2 - \frac{N}{2}zm,
\end{aligned}
\tag{5.15}
$$

where $\boldsymbol{x} = (x,y,z)$ and $J^{\mu}_{i_1,\ldots,i_{p/2}}, \eta_i, \theta_\mu \sim \mathcal{N}(0,1)$ are i.i.d. standard Gaussian variables.

Remark 11. An integral representation of the partition function of the original DAM model (5.11) is recovered by setting $t_0 = \frac{2\beta}{p!}$ and $(t,x,y,z) = (\frac{2\beta}{p!},0,0,0)$; then, by performing the Gaussian integration in (5.14), we recover the straight partition function.

In the case under consideration, the Boltzmann average induced by the interpolating partition function (5.14) reads as

$$
\omega_{t,\boldsymbol{x}}(O) = \frac{\sum_\sigma \int d\mu(\boldsymbol{\tau})\, O(\boldsymbol{\sigma},\boldsymbol{\tau}) B_{N,p,\xi,\alpha}(t,\boldsymbol{x})}{Z_{N,p,\xi,\alpha}(t,\boldsymbol{x})},
\tag{5.16}
$$

and, by averaging also with respect to the quenched noise,

$$
\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})\rangle_{t,\boldsymbol{x}} := \mathbb{E}_{\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\theta}}[\omega_{t,\boldsymbol{x}}(O(\boldsymbol{\sigma},\boldsymbol{\tau}))],
\tag{5.17}
$$

where $O(\boldsymbol{\sigma},\boldsymbol{\tau})$ is a generic observable in the configuration space of the system, $\mathbb{E}_{\boldsymbol{\xi},\boldsymbol{\eta},\boldsymbol{\theta}}$ denotes the quenched average over the set of i.i.d. Gaussian standard variables specified in the subscript, and $B_{N,p}$ is the generalized Boltzmann factor defined as

$$
\begin{aligned}
B_{N,p,\xi,\alpha}(t,\boldsymbol{x}) := \exp\Bigg(&\frac{tN}{2}m^p + \sqrt{\frac{t}{N^{p-1}}}\sqrt{\mathcal{V}}\sum_{\mu=2}^{K}\sum_{i_1<\cdots<i_{p/2}}^{N} J^{\mu}_{i_1,\ldots,i_{p/2}}\sigma_{i_1}\ldots\sigma_{i_{p/2}}\tau_\mu \\
&+ \sqrt{x}\sum_{i=1}^{N}\eta_i\sigma_i + \sqrt{N^{1-p/2}y}\sum_{\mu=1}^{K}\theta_\mu\tau_\mu - \frac{N^{1-p/2}}{2}[\mathcal{V}_0(t-t_0)+y]\sum_{\mu=1}^{K}\tau_\mu^2 + \frac{N}{2}zm\Bigg).
\end{aligned}
$$

As mentioned in Sec. I, the high-storage regime of associative neural networks exhibits both ferromagnetic and spin-glass features. Thus, in order to fully characterize the system behavior, besides the usual Mattis magnetizations, we need the overlap for the two sets of relevant variables in the integral formulation of the partition function (5.14).

Definition 12. The order parameters used to describe the macroscopic behavior of the model are the Mattis magnetization m *(already defined in (8) and used to quantify the retrieval capability of the network), the replica overlap in the $\boldsymbol{\sigma}$ variables,*

$$
q_{12} := \frac{1}{N}\sum_{i=1}^{N}\sigma_i^{(1)}\sigma_i^{(2)},
\tag{5.18}
$$

and the replica overlap in the $\boldsymbol{\tau}'$s variables,

$$
p_{12} := \frac{1}{N^{p/2}}\sum_{\mu=1}^{K}\tau_\mu^{(1)}\tau_\mu^{(2)}.
\tag{5.19}
$$

The main result of this section is given by the following theorem, which establishes that the expectation of the order parameters of the p-th order DAM in the high-storage limit fulfills a set of PDEs that generalize the Burgers hierarchy structure.

Theorem 3. *The high-storage regime for the DAM models under the RS assumption in the thermodynamic limit can be described by the following system of partial differential equation (PDEs):*

$$
\begin{cases}
\partial_t\bar{q} + \partial_x\bar{m}^p - \mathcal{V}_0\partial_x\bar{q}^{p/2}\bar{p} = 0,\\
\partial_t\bar{p} + \partial_y\bar{m}^p - \mathcal{V}_0\partial_y\bar{q}^{p/2}\bar{p} = 0,\\
\partial_t\bar{m} - \partial_z\bar{m}^p + \mathcal{V}_0\partial_z\bar{q}^{p/2}\bar{p} = 0,
\end{cases}
\tag{5.20}
$$

with the initial conditions

$$\begin{cases} \bar{q}(0,\boldsymbol{x}) = \mathbb{E}_\eta\left[\tanh^2\left(\sqrt{x}\eta + \frac{z}{2}\right)\right], \\ \bar{p}(0,\boldsymbol{x}) = \begin{cases} \frac{\alpha y}{(1+y-t_0)^2} & \textit{if } p = 2, \\ \alpha y & \textit{if } p = 2k \textit{ with } k = 2,3,\ldots, \end{cases} \\ \bar{m}(0,\boldsymbol{x}) = \mathbb{E}_\eta\left[\tanh\left(\sqrt{x}\eta + \frac{z}{2}\right)\right], \end{cases} \tag{5.21}$$

where $\mathbb{E}_\eta$ *is the Gaussian average over the variable* η*.*

Again, in order to prove the theorem, we need to undergo some preliminary passages.

Definition 13. For all even $p \geq 2$*, Guerra's action functional is defined as*

$$S_{N,p,\alpha}(t,\boldsymbol{x}) := 2A_{N,p,\alpha}(t,\boldsymbol{x}) - x. \tag{5.22}$$

Lemma 4. The partial derivatives of Guerra's action $S_{N,p}(t,\boldsymbol{x})$ *can be expressed in terms of the generalized expectations of the order parameters as*

$$\begin{aligned} \partial_t S_{N,p,\alpha} &= \langle m^p\rangle_{t,\boldsymbol{x}} - \mathcal{V}_0\langle p_{12}q_{12}^{p/2}\rangle_{t,\boldsymbol{x}}, \\ \partial_x S_{N,p,\alpha} &= -\langle q_{12}\rangle_{t,\boldsymbol{x}}, \\ \partial_y S_{N,p,\alpha} &= -\langle p_{12}\rangle_{t,\boldsymbol{x}}, \\ \partial_z S_{N,p,\alpha} &= \langle m\rangle_{t,\boldsymbol{x}}. \end{aligned} \tag{5.23}$$

The computation of the spacetime derivatives is fairly lengthy but straightforward. We report the computation of the derivatives in Appendix D.

In order to derive differential identities for the expectation values of the order parameters, we need to compute the spatial derivatives of a generic function of two replicas $O(\boldsymbol{\sigma}^{(1)},\boldsymbol{\sigma}^{(2)},\boldsymbol{\tau}^{(1)},\boldsymbol{\tau}^{(2)})$.

Lemma 5. Let $\underline{\boldsymbol{\sigma}} = (\boldsymbol{\sigma}^{(1)},\boldsymbol{\sigma}^{(2)})$ *and* $\underline{\boldsymbol{\tau}} = (\boldsymbol{\tau}^{(1)},\boldsymbol{\tau}^{(2)})$ *be the configurations of the 2-replicated system. Then,*

$$\begin{aligned} \partial_x\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})\rangle_{t,\boldsymbol{x}} = \frac{N}{2}\sum_{a,b=1}^{2}\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})q_{ab}\rangle_{t,\boldsymbol{x}} - 2N\sum_{a=1}^{2}\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})q_{a3}\rangle_{t,\boldsymbol{x}} - N\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})\rangle_{t,\boldsymbol{x}} \\ + 3N\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})q_{34}\rangle_{t,\boldsymbol{x}}, \end{aligned} \tag{5.24}$$

$$\begin{aligned} \partial_y\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})\rangle_{t,\boldsymbol{x}} = \frac{N}{2}\sum_{a,b=1}^{2}\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})p_{ab}\rangle_{t,\boldsymbol{x}} - 2N\sum_{a=1}^{2}\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})p_{a3}\rangle_{t,\boldsymbol{x}} - N\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})p_{33}\rangle_{t,\boldsymbol{x}} \\ + 3N\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})p_{34}\rangle_{t,\boldsymbol{x}}, \end{aligned} \tag{5.25}$$

and

$$\partial_z\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})\rangle_{t,\boldsymbol{x}} = N(\langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})m\rangle_{t,\boldsymbol{x}} - \langle O(\underline{\boldsymbol{\sigma}},\underline{\boldsymbol{\tau}})\rangle_{t,\boldsymbol{x}}\langle m\rangle_{t,\boldsymbol{x}}). \tag{5.26}$$

The complete proof is given in Appendix E.

Again, we will assume the RS in order to simplify the computations by neglecting the fluctuations of the order parameters with respect to their expectation values, in particular, the below proposition.

Proposition 3. The following equalities hold:

$$\langle m^p\rangle_{t,\boldsymbol{x}} = \left(\frac{1}{N}\partial_z + \langle m\rangle_{t,\boldsymbol{x}}\right)^{p-1}\langle m\rangle_{t,\boldsymbol{x}}, \tag{5.27}$$

$$\langle p_{12}q_{12}^{p/2}\rangle_{t,\boldsymbol{x}} = \left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,\boldsymbol{x}}\right)^{p/2}\langle p_{12}\rangle_{t,\boldsymbol{x}} + R_N^{(\frac{p}{2})}(t,\boldsymbol{x}), \tag{5.28}$$

where $R_N^{(\frac{p}{2})}(t,\boldsymbol{x})$ *collects the terms involving the fluctuations of the order parameters and thus vanishes in the* $N\to\infty$ *limit and under the RS assumption.*

Proof. To simplify the notation, we will drop the subscript $t, \boldsymbol{x}$ from the quenched averages. The derivation of (5.27) is achieved by iterating the property (5.26) with $O(\boldsymbol{\sigma}, \boldsymbol{\tau}) = m(\boldsymbol{\sigma})^{p-1}$. Let us now observe that

$$\langle p_{12}q_{12}^l q_{13}\rangle = \Big\langle p_{12}q_{12}^h \Delta(q_{13})\Big\rangle + \langle p_{12}q_{12}^l\rangle\langle q_{13}\rangle = \langle p_{12}q_{12}^l\rangle\langle q_{13}\rangle + R_N^{(1,l)}(t,\boldsymbol{x})$$

and

$$\langle p_{12}q_{12}^l q_{34}\rangle = \Big\langle p_{12}q_{12}^l \Delta(q_{34})\Big\rangle + \langle p_{12}q_{12}^l\rangle\langle q_{34}\rangle = \langle p_{12}q_{12}^l\rangle\langle q_{34}\rangle + R_N^{(2,l)}(t,\boldsymbol{x}),$$

where $\Delta(q_{ab}) := q_{ab} - \langle q_{ab}\rangle$ and $R_N^{(1,l)}, R_N^{(2,l)}$ collect the contributions involving the fluctuations.

Now, we use $O(\underline{\boldsymbol{\sigma}}, \underline{\boldsymbol{\tau}}) = p_{12}q_{12}^l$ in Lemma 5 and get

$$\begin{aligned}\partial_x\langle p_{12}q_{12}^l\rangle_{t,\boldsymbol{x}} &= \frac{N}{2}\Big(\langle p_{12}q_{12}^l\rangle_{t,\boldsymbol{x}} + \langle p_{12}q_{12}^l q_{12}\rangle_{t,\boldsymbol{x}} + \langle p_{12}q_{12}^l q_{21}\rangle_{t,\boldsymbol{x}} + \langle p_{12}q_{12}^l\rangle_{t,\boldsymbol{x}}\Big)\\ &\quad - 2N\Big(\langle p_{12}q_{12}^l q_{13}\rangle_{t,\boldsymbol{x}} + \langle p_{12}q_{12}^l q_{23}\rangle_{t,\boldsymbol{x}}\Big) - N\langle p_{12}q_{12}^l\rangle_{t,\boldsymbol{x}} + 3N\langle p_{12}q_{12}^l q_{34}\rangle_{t,\boldsymbol{x}}\\ &= N\langle p_{12}q_{12}^{l+1}\rangle_{t,\boldsymbol{x}} - 4N\langle p_{12}q_{12}^l q_{13}\rangle_{t,\boldsymbol{x}} + 3N\langle p_{12}q_{12}^l q_{34}\rangle_{t,\boldsymbol{x}},\end{aligned}$$

whence the following equality holds for all $l \in \mathbb{N}^+$:

$$\partial_x\langle p_{12}q_{12}^l\rangle_{t,\boldsymbol{x}} = N\Big(\langle p_{12}q_{12}^{l+1}\rangle_{t,\boldsymbol{x}} - 4\langle p_{12}q_{12}^l q_{13}\rangle_{t,\boldsymbol{x}} + 3\langle p_{12}q_{12}^l q_{34}\rangle_{t,\boldsymbol{x}}\Big). \tag{5.29}$$

Combining the previous expressions, we obtain

$$\begin{aligned}\langle p_{12}q_{12}^{l+1}\rangle &= \frac{1}{N}\partial_x\langle p_{12}q_{12}^l\rangle + 4\langle p_{12}q_{12}^l\rangle\langle q_{13}\rangle - 3\langle p_{12}q_{12}^l\rangle\langle q_{34}\rangle + R_N^{(l)}(t,\boldsymbol{x})\\ &= \frac{1}{N}\partial_x\langle p_{12}q_{12}^l\rangle + \langle p_{12}q_{12}^l\rangle\langle q_{12}\rangle + R_N^{(l)}(t,\boldsymbol{x}),\end{aligned}$$

where $R_N^{(l)}(t,\boldsymbol{x}) := 4R_N^{(1,l)}(t,\boldsymbol{x}) - 3R_N^{(2,l)}(t,\boldsymbol{x})$ and we used $\langle q_{34}\rangle = \langle q_{13}\rangle = \langle q_{12}\rangle$, which follows from invariance under replica labeling. The previous equation can be written as follows:

$$\langle p_{12}q_{12}^{l+1}\rangle = \left(\frac{1}{N}\partial_x + \langle q_{12}\rangle\right)\langle p_{12}q_{12}^l\rangle + R_N^{(l)}(t,\boldsymbol{x}). \tag{5.30}$$

Iterating the procedure, we get

$$\langle p_{12}q_{12}^{l+1}\rangle = \left(\frac{1}{N}\partial_x + \langle q_{12}\rangle\right)^{l+1}\langle p_{12}\rangle + R_N^{(\frac{p}{2})}(t,\boldsymbol{x}), \tag{5.31}$$

where $R_N^{(\frac{p}{2})}(t,\boldsymbol{x})$ collects all the terms involving the rest of previous expansions (and thus vanishes in the $N \to \infty$ limit and under the RS assumption). Then, by imposing $l = p/2 - 1$, we get the thesis. □

Now, we can use all the information obtained to build a PDE that can describe the thermodynamics of the DAM models. Indeed, recalling the temporal derivative of the Guerra's action (5.23) and using the result obtained in Proposition 3, we have

$$\partial_t S_{N,p,\alpha} = \left(\frac{1}{N}\partial_z + \langle m\rangle_{t,\boldsymbol{x}}\right)^{p-1}\langle m\rangle_{t,\boldsymbol{x}} - \mathcal{V}_0\left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,\boldsymbol{x}}\right)^{p/2}\langle p_{12}\rangle_{t,\boldsymbol{x}} - R_N^{(\frac{p}{2})}(t,\boldsymbol{x}). \tag{5.32}$$

Finally, taking the spatial derivatives of this expression and denoting $\boldsymbol{D}_N(t,\boldsymbol{x}) = -\nabla R_N^{(p/2)}$, we have

$$\begin{aligned}-\partial_t\langle q_{12}\rangle_{t,\boldsymbol{x}} - \partial_x\left(\frac{1}{N}\partial_z + \langle m\rangle_{t,\boldsymbol{x}}\right)^{p-1}\langle m\rangle_{t,\boldsymbol{x}} + \mathcal{V}_0\partial_x\left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,\boldsymbol{x}}\right)^{p/2}\langle p_{12}\rangle_{t,\boldsymbol{x}} &= D_{N,x},\\ -\partial_t\langle p_{12}\rangle_{t,\boldsymbol{x}} - \partial_y\left(\frac{1}{N}\partial_z - \langle m\rangle_{t,\boldsymbol{x}}\right)^{p-1}\langle m\rangle_{t,\boldsymbol{x}} + \mathcal{V}_0\partial_y\left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,\boldsymbol{x}}\right)^{p/2}\langle p_{12}\rangle_{t,\boldsymbol{x}} &= D_{N,y},\\ \partial_t\langle m\rangle_{t,\boldsymbol{x}} - \partial_z\left(\frac{1}{N}\partial_z + \langle m\rangle_{t,\boldsymbol{x}}\right)^{p-1}\langle m\rangle_{t,\boldsymbol{x}} + \mathcal{V}_0\partial_z\left(\frac{1}{N}\partial_x + \langle q_{12}\rangle_{t,\boldsymbol{x}}\right)^{p/2}\langle p_{12}\rangle_{t,\boldsymbol{x}} &= D_{N,z}.\end{aligned} \tag{5.33}$$

The lhs of the system of PDEs constitutes the 3 + 1-dimensional DAM generalization of the Burgers hierarchy structure. Similar to the Derrida model case, at finite N, we have a source term on the rhs, which vanishes in the limit $N \to \infty$ under the RS assumption of the order parameters. In this case, we can analyze the thermodynamic limit and describe the equilibrium dynamics of the model.

We are now ready to prove Theorem 3.

Proof 3. First, let us call

$$\bar{m}(t,\boldsymbol{x}) = \lim_{N\to\infty} \langle m\rangle_{t,\boldsymbol{x}}, \qquad \bar{q}(t,\boldsymbol{x}) = \lim_{N\to\infty} \langle q_{12}\rangle_{t,\boldsymbol{x}}, \qquad \bar{p}(t,\boldsymbol{x}) = \lim_{N\to\infty} \langle p_{12}\rangle_{t,\boldsymbol{x}}$$

the expectation values of the order parameters in the thermodynamic limit. Taking $N \to \infty$ in (5.33) and recalling that the source contributions $\boldsymbol{D}_N(t,\boldsymbol{x})$ vanish in this limit under the RS assumption, we arrive at the PDE system (5.20). Let us now find the initial conditions (5.21). To do this, we start calculating the interpolating partition function in $t = 0$,

$$\begin{aligned} Z_{N,p,\xi,\alpha}(0,\boldsymbol{x}) &= \sum_{\boldsymbol{\sigma}} \int d\mu(\boldsymbol{\tau}) \exp\left(\sqrt{x}\sum_{i=1}^N \eta_i\sigma_i + \sqrt{\frac{y}{N^{p/2-1}}}\sum_{\mu=1}^K \theta_\mu\tau_\mu - \frac{y-\mathcal{V}_0 t_0}{2N^{p/2-1}}\sum_{\mu=1}^K \tau_\mu^2 + \frac{z}{2}\sum_{i=1}^N \sigma_i\right) \\ &= \sum_{\boldsymbol{\sigma}} \exp\left(\sum_{i=1}^N (\sqrt{x}\eta_i + \frac{1}{2}z)\sigma_i\right) \int d\mu(\boldsymbol{\tau}) \exp\left(\sqrt{\frac{y}{N^{p/2-1}}}\sum_{\mu=1}^K \theta_\mu\tau_\mu - \frac{y-\mathcal{V}_0 t_0}{2N^{p/2-1}}\sum_{\mu=1}^K \tau_\mu^2\right) \\ &= \prod_{i=1}^N 2\cosh\left(\sqrt{x}\eta_i + \frac{z}{2}\right)\prod_{\mu=1}^K \frac{1}{\sqrt{N^{1-p/2}(y-\mathcal{V}_0 t_0)+1}} \exp\left(\frac{N^{1-p/2}\, y\theta_\mu^2}{2\left(N^{1-p/2}(y-\mathcal{V}_0 t_0)+1\right)}\right). \end{aligned}$$

By using the definition of the interpolating statistical pressure (2.4), we see that

$$\begin{aligned} A_{N,p,\alpha}(0,\boldsymbol{x}) &= \frac{1}{N}\sum_{i=1}^N \mathbb{E}_\eta \log 2\cosh\left(\sqrt{x}\eta_i + \frac{z}{2}\right) + \frac{1}{N}\sum_{\mu=1}^K \mathbb{E}_\theta\left(\frac{N^{1-p/2} y\theta_\mu^2}{2\left(N^{1-p/2}(y-\mathcal{V}_0 t_0)+1\right)}\right) \\ &\quad - \frac{K}{2N}\log\left(N^{1-p/2}(y-\mathcal{V}_0 t_0)+1\right) \\ &= \mathbb{E}_\eta \log 2\cosh\left(\sqrt{x}\eta + \frac{z}{2}\right) + \frac{K}{N^{p/2}}\frac{y}{2\left(N^{1-p/2}(y-\mathcal{V}_0 t_0)+1\right)} \\ &\quad - \frac{K}{2N}\log\left(N^{1-p/2}(y-\mathcal{V}_0 t_0)+1\right), \end{aligned}$$

where we used the fact that the η_i's are i.i.d. random variables and $\mathbb{E}_\theta[\theta_\mu^2] = 1$ for all $\mu = 1,\ldots,K$. Now, recalling Eq. (5.22), we can straightforwardly derive the initial condition for the order parameters according to (5.23). First,

$$\begin{aligned} q(0,\boldsymbol{x}) &= \lim_{N\to\infty}(-\partial_x S_{N,p,\alpha}(0,\boldsymbol{x})) = -\partial_x\left[\mathbb{E}_\eta \log 2\cosh\left(\sqrt{x}\eta + \frac{z}{2}\right) - x\right] \\ &= \mathbb{E}_\eta\left[\tanh^2\left(\sqrt{x}\eta + \frac{z}{2}\right)\right]. \end{aligned} \tag{5.34}$$

Analogously, we have

$$\begin{aligned} p(0,\boldsymbol{x}) &= \lim_{N\to\infty}(-\partial_y S_{N,p,\alpha}(0,\boldsymbol{x})) \\ &= -\lim_{N\to\infty}\partial_y\left(\frac{K}{N^{p/2}}\frac{y}{1+N^{1-p/2}(y-\mathcal{V}_0 t_0)} - \frac{K}{N}\log\left[1+N^{1-p/2}(y-\mathcal{V}_0 t_0)\right]\right) \\ &= -\lim_{N\to\infty}\left(-\frac{K}{N^{p-1}}\frac{y}{[1+N^{1-p/2}(y-\mathcal{V}_0 t_0)]^2}\right) \\ &= \begin{cases} \dfrac{\alpha y}{[1+(y-t_0)]^2} & \text{if } p=2, \\ \alpha y & \text{if } p=2k \ \ and \ \ k=2,3,\ldots, \end{cases} \end{aligned} \tag{5.35}$$

where we used the fact that for $p = 2$, we have $\mathcal{V}_0 = 1$. Finally,

$$m(0,\boldsymbol{x}) = \lim_{N\to\infty} \partial_z S_{N,p,\alpha}(0,\boldsymbol{x}) = \partial_z \mathbb{E}_\eta \log 2\cosh\left(\sqrt{x}\eta + \frac{z}{2}\right) = \mathbb{E}_\eta \tanh\left(\sqrt{x}\eta + \frac{z}{2}\right). \tag{5.36}$$

This concludes the proof. □

Corollary 3. The system of PDEs (5.20) can be rewritten in a nonlinear wave equation as

$$\partial_t \boldsymbol{\phi} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{\phi} = 0, \tag{5.37}$$

where $\boldsymbol{\phi} := (\bar{q},\bar{p},\bar{m})$ *is the vector of the order parameters and* $\boldsymbol{v} := \left(-\mathcal{V}_0 \frac{p}{2}\bar{q}^{p/2-1}\bar{p}, -\mathcal{V}_0\bar{q}^{p/2}, -p\bar{m}^{p-1}\right)$ *is the effective velocity.*

Proof. We prove Eq. (5.37) for the first component, as the others follow accordingly. Let us define the function $G(\boldsymbol{\phi}) = \bar{m}^p - \mathcal{V}_0\bar{q}^{p/2}\bar{p}$, so that the PDE for the order parameter $\bar{q}$ can be rewritten as

$$\partial_t\bar{q} + \partial_x G(\boldsymbol{\phi}) = 0.$$

The x-derivative of the function G is straightforwardly computed,

$$\partial_x G(\boldsymbol{\phi}) = \partial_x(\bar{m}^p - \mathcal{V}_0\bar{q}^{p/2}\bar{p}) = p\bar{m}^{p-1}\partial_x\bar{m} - \frac{p}{2}\mathcal{V}_0\bar{q}^{p/2-1}\bar{p}\,\partial_x\bar{q} - \mathcal{V}_0\bar{q}^{p/2}\partial_x\bar{p}.$$

Now,

$$\partial_x\bar{p} = \partial_x\left(-\lim_{N\to\infty}\partial_y S_{N,p,\alpha}(t,\boldsymbol{x})\right) = \partial_y\left(-\lim_{N\to\infty}\partial_x S_{N,p,\alpha}(t,\boldsymbol{x})\right) = \partial_y\bar{q}.$$

In the same way,

$$\partial_x\bar{m} = \partial_x\left(\lim_{N\to\infty}\partial_z S_{N,p,\alpha}(t,\boldsymbol{x})\right) = \partial_z\left(\lim_{N\to\infty}\partial_x S_{N,p,\alpha}(t,\boldsymbol{x})\right) = -\partial_z\bar{q}.$$

With these results, we have

$$\partial_x G(\boldsymbol{\phi}) = -\frac{p}{2}\mathcal{V}_0\bar{q}^{p/2-1}\bar{p}\,\partial_x\bar{q} - \mathcal{V}_0\bar{q}^{p/2}\partial_y\bar{q} - p\bar{m}^{p-1}\partial_z\bar{q} = (v_x\partial_x + v_y\partial_y + v_z\partial_z)\bar{q} = (\boldsymbol{v}\cdot\nabla)\bar{q}.$$

This leads to

$$\partial_t\bar{q} + (\boldsymbol{v}\cdot\nabla)\bar{q} = 0.$$

□

Corollary 4. The equilibrium dynamics of the DAM models is given by the following set of self-consistency equations:

$$\begin{aligned}
\bar{q} &= \mathbb{E}_\eta\left[\tanh^2\left(\sqrt{\beta'\frac{p}{2}\mathcal{V}_0\bar{q}^{p/2-1}\bar{p}}\,\eta + \beta'\frac{p}{2}\bar{m}^{p-1}\right)\right],\\
\bar{m} &= \mathbb{E}_\eta\left[\tanh\left(\sqrt{\beta'\frac{p}{2}\mathcal{V}_0\bar{q}^{p/2-1}\bar{p}}\,\eta + \beta'\frac{p}{2}\bar{m}^{p-1}\right)\right],\\
\bar{p} &= \begin{cases} \dfrac{\alpha\beta\bar{q}}{[1-\beta(1-\bar{q})]^2} & \text{if } p = 2,\\ \alpha\beta'\mathcal{V}_0\bar{q}^{p/2} & \text{if } p = 2k \text{ with } k = 2,3,\ldots, \end{cases}
\end{aligned} \tag{5.38}$$

where $\beta' := \frac{2\beta}{p!}$.

Proof. In order to prove our assertion, we use the vector PDE (5.37), whose solution can be given in implicit form as

$$\boldsymbol{\phi}(t,\boldsymbol{x}) = \boldsymbol{\phi}_0(\boldsymbol{x} - \boldsymbol{v}t),$$

where $\boldsymbol{\phi}_0(\boldsymbol{x})$ is the initial profile given by conditions (5.21). For the first component, we have

$$\begin{aligned}
\bar{q}(t,\boldsymbol{x}) = \phi_{0,x}(\boldsymbol{x} - \boldsymbol{v}t) &= \mathbb{E}_\eta \tanh^2\left(\sqrt{x - v_x t}\,\eta + \frac{z - v_z t}{2}\right)\\
&= \mathbb{E}_\eta \tanh^2\left(\sqrt{x + \frac{p}{2}\mathcal{V}_0\bar{q}^{p/2-1}\bar{p}t}\,\eta + \frac{z + p\bar{m}^{p-1}t}{2}\right).
\end{aligned} \tag{5.39}$$

Analogously,

$$\bar{m}(t,\boldsymbol{x}) = \phi_{0,z}(\boldsymbol{x}-\boldsymbol{v}t) = \mathbb{E}_\eta \tanh\left(\sqrt{x+\frac{p}{2}\mathcal{V}_0\bar{q}^{p/2-1}\bar{p}t}\,\eta + \frac{z+p\bar{m}^{p-1}t}{2}\right). \tag{5.40}$$

Finally, if $p = 2$, we have

$$\bar{p}(t,\boldsymbol{x}) = \phi_{0,y}(\boldsymbol{x}-\boldsymbol{v}t) = \frac{\alpha(y-v_y t)}{(1+y-v_y t-t_0)^2} = \frac{\alpha(y+\bar{q}t)}{(1+y+\bar{q}t-t_0)^2}, \tag{5.41}$$

while, for $p = 2k$ with $k \geq 2$, the same order parameter satisfies the self-consistency equation

$$\bar{p}(t,\boldsymbol{x}) = \phi_{0,y}(\boldsymbol{x}-\boldsymbol{v}t) = \alpha(y-v_y t) = \alpha(y+\mathcal{V}_0\bar{q}^{p/2}t). \tag{5.42}$$

Recalling that the thermodynamics of the DAM models is reproduced when $t = t_0 = \beta' = \frac{2\beta}{p!}$ and $\boldsymbol{x} = \boldsymbol{0}$, we easily get the thesis. □

Remark 12. The self-consistency equations in Corollary 4 are in agreement with those obtained by Gardner in Ref. 53 by means of the replica approach.

Remark 13. As expected, for $p = 2$, the expressions given in Eq. (5.38) reduce to the Hopfield case. For $p = 2k > 2$, upon eliminating $\bar{p}$ from the self-consistency equations, we get

$$\bar{m} = \mathbb{E}_\eta \tanh\left[\frac{\beta}{(p-1)!}\left(\sqrt{\frac{2\alpha\mathcal{V}_0^2}{p}\bar{q}^{p-1}}\,\eta + \bar{m}^{p-1}\right)\right], \tag{5.43}$$

$$\bar{q} = \mathbb{E}_\eta \tanh^2\left[\frac{\beta}{(p-1)!}\left(\sqrt{\frac{2\alpha\mathcal{V}_0^2}{p}\bar{q}^{p-1}}\,\eta + \bar{m}^{p-1}\right)\right], \tag{5.44}$$

suggesting that the system has an effective thermal noise level of $\beta_{\rm eff} = \beta/(p-1)!$ and an effective capacity $\alpha_{\rm eff} = 2\alpha\mathcal{V}_0^2/p$ (controlling the level of glassy noise in the model). For large enough p, we have $(2p-1)!! \gg ((p-1)!!)^2$, so that

$$\frac{2\mathcal{V}_0^2}{p} = \frac{[(2p-1)!!-((p-1)!!)^2]}{p}) \simeq \frac{(2p-1)!!}{p} = \frac{(2p-1)(2p-3)!!}{p} \simeq 2(2p-3)!!,$$

meaning that

$$\frac{K}{N^{p-1}} \sim \frac{\alpha_{\rm eff}}{2(2p-3)!!},$$

which is the same scaling found by Hopfield and Krotov[19] in the perfect retrieval definition of the critical storage capacity.

By solving these self-consistency Eq. (5.38) numerically by a fixed-point method for a given p and tuning the parameters T and α, we obtain the phase diagrams shown in Fig. 2. As expected, the diagrams exhibit the existence of three different regions:

- For high levels of noise T, no matter the value of storage α, the only stable solution is given by $\bar{m} = 0$, $\bar{q} = 0$, thus the system is ergodic (III).
- At lower temperatures and with relatively high load, the system exhibits spin-glass behaviors (II), and the solution is characterized by $\bar{m} = 0$ and $\bar{q} \neq 0$.
- For relatively small values of α and T, we have $\bar{m}, \bar{q} \neq 0$ and the system is located in the retrieval phase (I). In this situation, the system behaves as an associative neural network performing pattern recognition spontaneously. In particular, we can see that the retrieval region observed for values of T and α relatively small can be further split in a pure retrieval region—where pure states are global minima for the free energy—and a mixed retrieval region—where pure states are local minima, yet their attraction basin is large enough for the system to end there if properly stimulated.

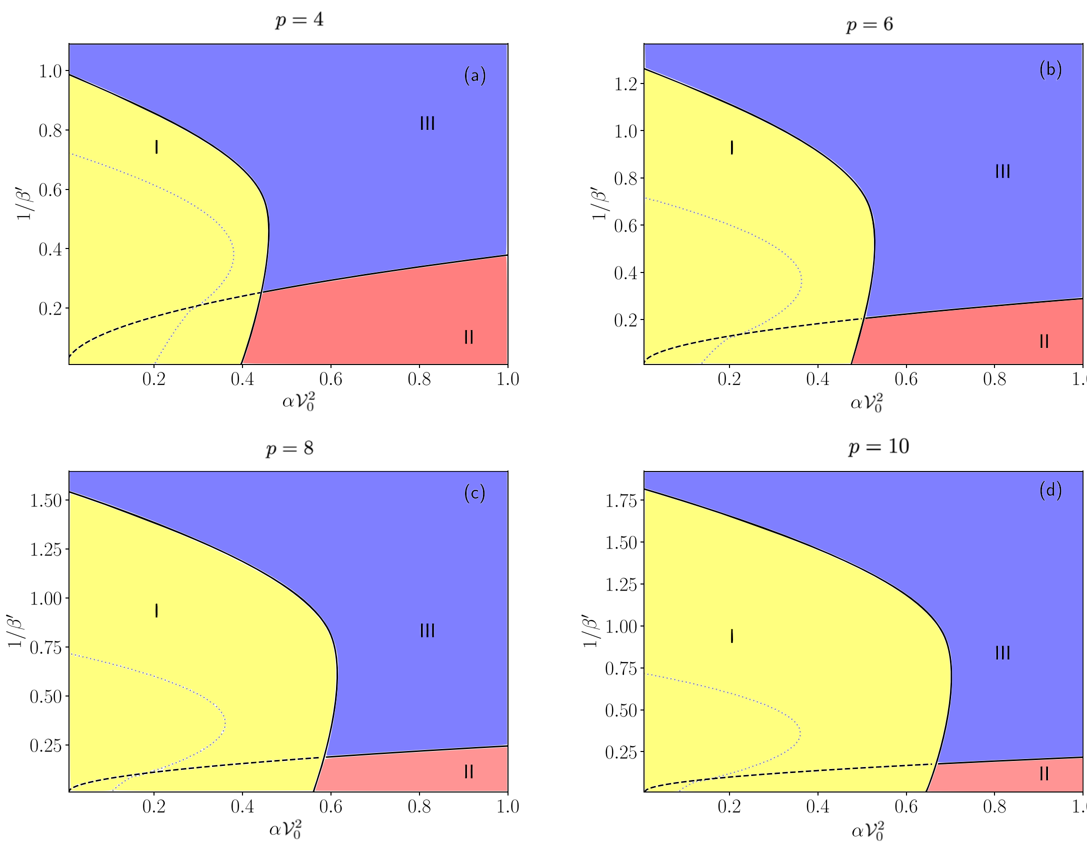

FIG. 2. Phase diagram for p = 4 [panel (a)], p = 6 [panel (b)], p = 8 [panel (c)], and p = 10 [panel (d)], in the space of control parameters α [through the combination $\alpha\mathcal{V}_0^2$, with the latter factor given by Eq. (5.13)] and $\beta' = 2\beta/p!$. Region I (in yellow), delimited by the solid black line, is the retrieval region, while region II (in red) and region III (in blue) are, respectively, the spin-glass and the ergodic phases. The dashed line is the prolongation of the spin-glass phase inside the retrieval region. Finally, the lighter dotted line inside region I identifies the portion of the parameter space in which the retrieval states are global minima for the free energy. Notice that the indentation that can be seen in the transition line delimiting the retrieval phase is a spurious effect due to the RS approximation.[48]

Thus, by increasing p, we need to afford higher costs in terms of resources since the number of connection weights to be properly set grows as $\binom{N}{p}$, but we also have a reward on a coarse scale, since the number of storable patterns grows as $K \sim N^{p-1}$, as well as on a fine scale, since the critical load α_c also increases with p.

VI. CONCLUSIONS

In this work, we focused on DAMs, which are neural-network models widely used for pattern recognition tasks and characterized by high-order (higher than quadratic) interactions between the constituting neurons. Extensive empirical evidence has shown that these models significantly outperform non-dense networks (displaying only quadratic interactions), especially as for the ability to correctly recognize adversarial or extremely noisy examples,[19–21,49] hence making these models particularly suitable for detecting and coping with malicious attacks. From the theoretical side, results are sparse and mainly based on the possibility of recasting these networks as spin-glass-like models with spins interacting p-wise; these models can, in turn, be effectively handled by tools stemming from the statistical mechanics of disordered systems (e.g., 44 and 54). Here, we pave this way and develop analytical techniques for their investigation. More precisely, we translate the original statistical-mechanical problem into an analytical-mechanical one where control parameters play the role of spacetime coordinates, the free energy plays the role of action, and the macroscopic observables that assess whether the system can be used for pattern recognition tasks play the role of effective velocities and are shown to fulfill a set of nonlinear partial differential equations. In this framework, transitions from different regimes (e.g., from a region in the control parameter space where the system performs correctly and another one where information processing capabilities are lost) appear as the emergence of shock waves.

A main advantage of this route is that it allows for rigorous investigations in a field where most knowledge is based on (pseudo) heuristic approaches, with a wide set of already available methods and strategies to rely upon. Furthermore, by bridging two different perspectives, statistical mechanics and analytical methods, we anticipate a cross-fertilization that may lead to a deeper comprehension of the system's subtle mechanisms and ultimately advance the development of a complete theory for (deep) machine learning.

ACKNOWLEDGMENTS

E.A. and A.F. would like to thank A. Moro and F. Alemanno for useful discussions. The authors acknowledge the Sapienza University of Rome for financial support (Grant Nos. RM120172B8066CB0, RM12117A8590B3FA, and AR12117A623B0114).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Elena Agliari: Conceptualization (equal); Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal). **Alberto Fachechi**: Conceptualization (equal); Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal). **Chiara Marullo**: Conceptualization (equal); Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: PROOF OF LEMMA 1

Proof. First of all, we compute the following temporal derivative:

$$\begin{aligned}\partial_t A_{N,p}(t,x) &= \frac{1}{N}\mathbb{E}_{\boldsymbol{J}}\frac{1}{Z_{N,p,\boldsymbol{J}}(t,x)}\sum_{\boldsymbol{\sigma}}\sqrt{\frac{p!}{2N^{p-1}}}\frac{1}{2\sqrt{t}}\sum_{1\le i_1<\cdots<i_p\le N} J_{i_1\ldots i_p}\sigma_{i_1}\ldots\sigma_{i_p}B_{N,p,\boldsymbol{J}}(t,x)\\ &= \frac{1}{N}\sqrt{\frac{p!}{2N^{p-1}}}\frac{1}{2\sqrt{t}}\sum_{1\le i_1<\cdots<i_p\le N}\mathbb{E}_{\boldsymbol{J}}J_{i_1\ldots i_p}\omega_{t,x}(\sigma_{i_1}\ldots\sigma_{i_p}).\end{aligned}\tag{A1}$$

Here, we can use the Wick–Isserlis theorem for normally distributed random variables, ensuring that $\mathbb{E}_{\boldsymbol{J}}J_l f(\boldsymbol{J}) = \mathbb{E}_{\boldsymbol{J}}\partial_{J_l} f(\boldsymbol{J})$ for each function f of the quenched disorder $\boldsymbol{J}$. Thus,

$$\begin{aligned}\partial_t A_{N,p}(t,x) &= \frac{1}{N}\sqrt{\frac{p!}{2N^{p-1}}}\frac{1}{2\sqrt{t}}\sum_{1\le i_1<\cdots<i_p\le N}\mathbb{E}_{\boldsymbol{J}}\partial_{J_{i_1\ldots i_p}}\omega_{t,x}(\sigma_{i_1}\ldots\sigma_{i_p})\\ &= \frac{1}{N}\sqrt{\frac{p!}{2N^{p-1}}}\frac{1}{2\sqrt{t}}\sqrt{\frac{tp!}{2N^{p-1}}}\sum_{1\le i_1<\cdots<i_p\le N}\mathbb{E}_{\boldsymbol{J}}(\omega_{t,x}(1)-\omega_{t,x}(\sigma_{i_1}\ldots\sigma_{i_p})^2)\\ &= \frac{p!}{4N^p}\sum_{1\le i_1<\cdots<i_p\le N}(1-\mathbb{E}_{\boldsymbol{J}}\omega_{t,x}(\sigma_{i_1}\ldots\sigma_{i_p})^2).\end{aligned}\tag{A2}$$

The nontrivial contribution in round brackets in the last equality can be expressed in terms of the overlap order parameter. Indeed,

$$\begin{aligned}\mathbb{E}_{\boldsymbol{J}}\omega_{t,x}(\sigma_{i_1}\ldots\sigma_{i_p})^2 &= \mathbb{E}_{\boldsymbol{J}}\omega^{(1)}_{t,x}(\sigma^{(1)}_{i_1}\ldots\sigma^{(1)}_{i_p})\ \omega^{(2)}_{t,x}(\sigma^{(2)}_{i_1}\ldots\sigma^{(2)}_{i_p})\\ &= \mathbb{E}_{\boldsymbol{J}}\Omega^{(2)}_{t,x}(\sigma^{(1)}_{i_1}\sigma^{(2)}_{i_1}\ldots\sigma^{(1)}_{i_p}\sigma^{(2)}_{i_p})\\ &= \langle\sigma^{(1)}_{i_1}\sigma^{(2)}_{i_1}\ldots\sigma^{(1)}_{i_p}\sigma^{(2)}_{i_p}\rangle_{t,x},\end{aligned}\tag{A3}$$

where 1, 2 are the replica indices. Now, using Remark 2, in the thermodynamic limit, the following equality holds:

$$\partial_t A_{N,p}(t,x) = \frac{1}{4N^p}\sum_{i_1,\ldots,i_p=1}^{N}(1-\langle\sigma_{i_1}^{(1)}\sigma_{i_1}^{(2)}\ldots\sigma_{i_p}^{(1)}\sigma_{i_p}^{(2)}\rangle_{t,x})$$
$$= \frac{1}{4}\left(1-\left\langle\left(\frac{1}{N}\sum_{i=1}^{N}\sigma_i^{(1)}\sigma_i^{(2)}\right)^p\right\rangle_{t,x}\right) = \frac{1}{4}(1-\langle q_{12}^p\rangle_{t,x}). \tag{A4}$$

Concerning the spatial derivative, we proceed in the same way,

$$\partial_x A_{N,p}(t,x) = \frac{1}{N}\mathbb{E}_{\boldsymbol{J}}\frac{1}{Z_{\mathrm{N,p},\boldsymbol{J}}(t,x)}\sum_{\boldsymbol{\sigma}}\frac{1}{2\sqrt{x}}\sum_{i=1}^{N}J_i\sigma_i B_{N,p,\boldsymbol{J}}(t,x)$$
$$= \frac{1}{2N\sqrt{x}}\sum_{i=1}^{N}\mathbb{E}_{\boldsymbol{J}}J_i\omega_{t,x}(\sigma_i) = \frac{1}{2N\sqrt{x}}\sum_{i=1}^{N}\mathbb{E}_{\boldsymbol{J}}\partial_{J_i}\omega_{t,x}(\sigma_i)$$
$$= \frac{1}{2N}\sum_{i=1}^{N}(1-\mathbb{E}_{\boldsymbol{J}}\omega_{t,x}(\sigma_i)^2). \tag{A5}$$

In this case, we express $\mathbb{E}_{\boldsymbol{J}}\omega_{t,x}(\sigma_i)^2 = \langle\sigma_i^{(1)}\sigma_i^{(2)}\rangle_{t,x}$. Thus,

$$\partial_x A_{N,p}(t,x) = \frac{1}{2}(1-\langle q_{12}\rangle_{t,x}). \tag{A6}$$

By simply exploiting Definition 6 and the Remark 2, we get the thesis. □

APPENDIX B: PARTICULAR CASES OF LOW-STORAGE DAMS

In this appendix, having clarified the equations describing the general case of the DAM models in the low-storage regime, we will study two special cases: the standard case where $p = 2$ and the more complex case with $p = 3$. In particular, we will observe that these two cases can be described by the Burgers and Sharma–Tasso–Olver equations in a $(K+1)$-dimensional space, respectively. To start, however, we first need the following definition and lemma.

Lemma 6. For all $\mu, \nu = 1, \ldots, K$, we have

1. $[D_\mu, D_\nu] = 0$ *and*
2. $[\partial_\mu, D_\nu^s] = N[D_\nu^s, u_\mu]$ $\forall s > 0$,

where $[\cdot,\cdot]$ is the usual commutator.

Proof. The proof of statement 1. works by direct computation. Indeed,

$$[D_\mu, D_\nu] = \left[\frac{1}{N}\partial_\mu + u_\mu, \frac{1}{N}\partial_\nu + u_\nu\right]$$
$$= \frac{1}{N^2}[\partial_\mu,\partial_\nu] + \frac{1}{N}[\partial_\mu, u_\nu] + \frac{1}{N}[u_\mu,\partial_\nu] + [u_\mu, u_\nu]$$
$$= \frac{1}{N}(\partial_\mu u_\nu - \partial_\nu u_\mu).$$

Since the field $u_\mu(t,\boldsymbol{x})$ is conservative, i.e., $u_\mu(t,\boldsymbol{x}) = \partial_\mu A_{N,p,\xi,K}(t,\boldsymbol{x})$, we have

$$\partial_\mu u_\nu - \partial_\nu u_\mu = (\partial_\mu\partial_\nu - \partial_\nu\partial_\mu)A_{N,p,\xi,K}(t,\boldsymbol{x}) = 0,$$

meaning that $[D_\mu, D_\nu] = 0$. Let as now prove property 2. In this case, the proof works by exploiting the property $[D_\mu, D_\nu] = 0$ (from which it follows that $[D_\mu, D_\nu^s] = 0$) and the definition $D_\mu = \frac{1}{N}\partial_\mu + u_\mu$. Indeed,

$$0 = [D_\mu, D_\nu^s] = \frac{1}{N}[\partial_\mu, D_\nu^s] + [u_\mu, D_\nu^s].$$

By rearranging the equality, we easily get the thesis. □

From the previous lemma, we can then prove the following two propositions.

Proposition 4. In the $p = 2$ case with a generic finite K, the evolutive equation (5.10) reduces to the multidimensional Burgers equation.

Proof. In the $p = 2$ case, Eq. (5.10) reduces to

$$\partial_t u_\mu = -\sum_{\nu=1}^{K} \partial_\mu D_\nu u_\nu.$$

Now, using the second claim of Lemma 6 with $s = 1$, we have

$$\partial_\mu D_\nu u_\nu = (D_\nu \partial_\mu + N[D_\nu, u_\mu])u_\nu.$$

Recalling the definition of the D operator, we have

$$D_\nu \partial_\mu u_\nu = \left(\frac{1}{N}\partial_\nu + u_\nu\right)\partial_\mu u_\nu = \frac{1}{N}\partial_\nu \partial_\mu u_\nu + u_\nu \partial_\mu u_\nu,$$
$$[D_\nu, u_\mu] = \left[\frac{1}{N}\partial_\nu + u_\nu, u_\mu\right] = \frac{1}{N}[\partial_\nu, u_\mu] = \frac{1}{N}\partial_\nu u_\mu.$$

Thus,

$$\partial_t u_\mu = -\sum_{\nu=1}^{K}\left(\frac{1}{N}\partial_\nu \partial_\mu u_\nu + u_\nu \partial_\mu u_\nu + u_\nu \partial_\nu u_\mu\right).$$

However, now,

$$\partial_\nu \partial_\mu u_\nu = \partial_\nu \partial_\mu \partial_\nu A_{N,p=2,\xi,K}(t,\boldsymbol{x}) = \partial_\nu^2 u_\mu,$$
$$u_\nu \partial_\mu u_\nu = u_\nu \partial_\mu \partial_\nu A_{N,p=2,\xi,K}(t,\boldsymbol{x}) = u_\nu \partial_\nu u_\mu.$$

Using these results, we can rewrite the equation as

$$\partial_t u_\mu + \sum_{\nu=1}^{K}\left(\frac{1}{N}\partial_\nu^2 u_\mu + 2u_\nu \partial_\nu u_\mu\right) = 0$$

or, in vector form,

$$\partial_t \boldsymbol{u} + \frac{1}{N}\nabla^2 \boldsymbol{u} + 2(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = 0,$$

which is precisely the Burgers equation in $K + 1$ spacetime. □

Proposition 5. In the $p = 3$ case (and generic K), the evolutive equation (5.10) reduces to the multidimensional Sharma–Tasso–Olver (STO) equation.[55,56]

Proof. In the $p = 3$ case, Eq. (5.10) reduces to

$$\partial_t u_\mu = -\sum_{\nu=1}^{K} \partial_\mu D_\nu^2 u_\nu.$$

Recalling that $D_\nu := \left(\frac{1}{N}\partial_\nu + u_\nu\right)$, we have

$$\begin{aligned}\partial_\mu D_\nu^2 u_\nu &= \partial_\mu\left(\frac{1}{N}\partial_\nu + u_\nu\right)\left(\frac{1}{N}\partial_\nu + u_\nu\right)u_\nu \\ &= \partial_\mu\left[\frac{1}{N^2}\partial_\nu^2 u_\nu + \frac{1}{N}\partial_\nu(u_\nu^2) + \frac{1}{N}u_\nu\partial_\nu u_\nu + u_\nu^3\right] \\ &= \partial_\mu\left[\frac{1}{N^2}\partial_\nu^2 u_\nu + \frac{3}{N}u_\nu\partial_\nu u_\nu + u_\nu^3\right].\end{aligned}$$

Performing the derivative with respect to the μ-th component, we, therefore, get

$$\partial_\mu D_\nu^2 u_\nu = \frac{1}{N^2}\partial_\mu \partial_\nu^2 u_\nu + \frac{3}{N}(\partial_\mu u_\nu \partial_\nu u_\nu + u_\nu \partial_\mu \partial_\nu u_\nu) + 3u_\nu^2 \partial_\mu u_\nu. \tag{B1}$$

Now, recalling that $u_\mu(t,\boldsymbol{x}) := \omega_{t,\boldsymbol{x}}(m_\nu(\boldsymbol{\sigma}))$ and (5.7), we have

$$\begin{aligned}\partial_\mu\partial_\nu^2 u_\nu &= \partial_\mu\partial_\nu^3 A_{N,p=3,\xi,K}(t,\boldsymbol{x}) = \partial_\nu^3 u_\mu,\\ \partial_\mu u_\nu &= \partial_\mu\partial_\nu A_{N,p=3,\xi,K}(t,\boldsymbol{x}) = \partial_\nu u_\mu,\\ \partial_\mu\partial_\nu u_\nu &= \partial_\mu\partial_\nu^2 A_{N,p=3,\xi,K}(t,\boldsymbol{x}) = \partial_\nu^2 u_\mu.\end{aligned}$$

Using these results, we can rewrite the equation as

$$\partial_t u_\mu = -\sum_{\nu=1}^{K}\left(\frac{1}{N^2}\partial_\nu^3 u_\mu + \frac{3}{N}(\partial_\nu u_\mu)^2 + \frac{3}{N}u_\nu\partial_\nu^2 u_\mu + 3u_\nu^2\partial_\nu u_\mu\right).$$

□

APPENDIX C: UNIVERSALITY OF NOISE IN DAMS

In this appendix, we show how quenched-noise contributions in the partition function (5.11) can be described in terms of Gaussian variables, so that we can take benefit of Wick's theorem. Let us start by considering the noisy contribution generated by non-retrieved patterns,

$$\frac{1}{N^{p-1}}\sum_{\mu\geq 2}\sum_{i_1,\ldots,i_p}\xi_{i_1}^\mu\ldots\xi_{i_p}^\mu\sigma_{i_1}\ldots\sigma_{i_p} = N\sum_{\mu\geq 2}\left(\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu\sigma_i\right)^p \equiv N\sum_{\mu\geq 2}m_\mu^p,$$

where m_μ are the Mattis magnetizations. Under the assumption of single-pattern retrieval of $\boldsymbol{\xi}^1$, it is easy to show that $m_\mu \sim \mathcal{N}(0,N^{-1/2})$ for $\mu\geq 2$. With a straightforward application of Cramer's lemma for large deviation theory (see, for example, Ref. 57), it can be shown that the variable $s_{\mu,l}\equiv m_\mu^l$ is described by the probability distribution

$$P(s_{\mu,l}) \propto s_{\mu,p}^{-(1-\frac{1}{l})}\ \exp\left(-\frac{N}{2}s_{\mu,l}^{\frac{2}{l}}\right).$$

Clearly, for $l = 1$, this reduces to a Gaussian distribution as expected. Conversely, for $l > 2$, the probability distribution has support on $]0,+\infty[$, with first and second moments given by

$$\begin{aligned}\mathbb{E}s_{\mu,l} &= \frac{2^{\frac{l}{2}}N^{-\frac{l}{2}}\Gamma\left(\frac{l}{2}+\frac{1}{2}\right)}{\sqrt{\pi}},\\ \mathbb{E}s_{\mu,l}^2 &= \frac{2^l N^{-l}\Gamma\left(l+\frac{1}{2}\right)}{\sqrt{\pi}}.\end{aligned}$$

In particular, for $l = p$ even, we have $\mathbb{E}s_{\mu,p} = N^{-\frac{p}{2}}(p-1)!!$ and $\mathbb{E}s_{\mu,p}^2 = N^{-p}(2p-1)!!$ so that

$$\mathrm{Var}(s_{\mu,p}) = N^{-p}[(2p-1)!! - ((p-1)!!)^2].$$

Since the $s_{\mu,p}$'s are i.i.d. and their variance is finite, the sum

$$\frac{1}{N^p}\sum_{\mu\geq 2}\sum_{i_1,\ldots,i_p}\xi_{i_1}^\mu\ldots\xi_{i_p}^\mu\sigma_{i_1}\ldots\sigma_{i_p} = \sum_{\mu\geq 2}s_{\mu,p}$$

converges to a Gaussian distributed variable by virtue of the Central Limit Theorem (CLT), in particular,

$$\frac{1}{N^p}\sum_{\mu\geq 2}\sum_{i_1,\ldots,i_p}\xi_{i_1}^\mu\ldots\xi_{i_p}^\mu\sigma_{i_1}\ldots\sigma_{i_p}\xrightarrow[d]{}\mathcal{N}(\mu,\sigma^2)$$

for $K\to\infty$, with

$$\mu = N^{-\frac{p}{2}}K(p-1)!!, \tag{C1}$$

$$\sigma = \sqrt{KN^{-p}[(2p-1)!! - ((p-1)!!)^2]}. \tag{C2}$$

This means that the noise contribution $N\sum_{\mu\geq 2}m_\mu^p \sim N\mu + zN\sigma$, where $z\sim\mathcal{N}(0,1)$. Notice that the contribution of the expectation value can be neglected from the whole partition function as it would contribute to the free energy with a factor of order $\mathcal{O}(\log N/N)$.

In the spirit of Gaussian universality of noise generated by non-retrieved patterns,[50–52] we consider the random variable

$$N\tilde{\mathcal{V}}\sum_{\mu\geq 2}\left(\sum_{i_1<\cdots<i_{p/2}} J^{\mu}_{i_1,\ldots,i_{p/2}}\sigma_{i_1}\ldots\sigma_{i_{p/2}}\right)^2, \tag{C3}$$

with $J^{\mu}_{i_1,\ldots,i_{p/2}} \sim_{iid} \mathcal{N}(0,1)$ and $\tilde{\mathcal{V}}$ a free parameter to be suitably tuned. By carrying out a totally analogous analysis as before, it is easy to show that, in the large N and K limit,

$$\tilde{\mathcal{V}}\sum_{\mu\geq 2}\left(\sum_{i_1<\cdots<i_{p/2}} J^{\mu}_{i_1,\ldots,i_{p/2}}\sigma_{i_1}\ldots\sigma_{i_{p/2}}\right)^2 \xrightarrow[d]{} \mathcal{N}(\tilde{\mu},\tilde{\sigma}^2),$$

where $\tilde{\mu} = \tilde{\mathcal{V}}K\binom{N}{p/2}$ and $\tilde{\sigma} = \tilde{\mathcal{V}}\sqrt{2K}\binom{N}{p/2}$. Thus, the random variable (C3) behaves, in the thermodynamic limit, as $N\tilde{\mu} + zN\tilde{\sigma}$, with $z \sim \mathcal{N}(0,1)$. Again, we can forget about the first moment, as it would contribute as $\mathcal{O}(\log N/N)$ to the free energy, and thus tune the $\tilde{\mathcal{V}}$ to match the second moment (C2) of the noise generated by non-retrieved patterns,

$$\sigma = \tilde{\sigma} \Longrightarrow \sqrt{KN^{-p}[(2p-1)!! - ((p-1)!!)^2]} = \tilde{\mathcal{V}}\sqrt{2K}\binom{N}{p/2}, \tag{C4}$$

leading to

$$\tilde{\mathcal{V}} = \frac{N^{-\frac{p}{2}}}{\binom{N}{p/2}}\sqrt{\frac{(2p-1)!! - ((p-1)!!)^2]}{2}}.$$

In the large N limit, we have

$$\binom{N}{p/2} \sim N^{\frac{p}{2}}\left(\frac{p}{2}!\right)^{-1}.$$

Thus,

$$\tilde{\mathcal{V}} = N^{-p}\frac{p}{2}!\sqrt{\frac{(2p-1)!! - ((p-1)!!)^2]}{2}} = N^{-p}\mathcal{V}.$$

To conclude, we see that, in the $N, K \to \infty$ limit,

$$\frac{1}{N^{p-1}}\sum_{\mu\geq 2}\sum_{i_1,\ldots,i_p}\xi^{\mu}_{i_1}\ldots\xi^{\mu}_{i_p}\sigma_{i_1}\ldots\sigma_{i_p} \sim \frac{1}{N^{p-1}}\mathcal{V}\sum_{\mu\geq 2}\left(\sum_{i_1<\ldots<i_{p/2}} J^{\mu}_{i_1,\ldots,i_{p/2}}\sigma_{i_1}\ldots\sigma_{i_{p/2}}\right)^2,$$

where, here, the symbol ~ means that the two sides of the equation are two Gaussian variables with the same second moment (which is the only relevant momentum for our concerns).

APPENDIX D: PROOF OF LEMMA 4

Proof. We prove the equality for the t-derivative of the Guerra's action, as the others follow from similar calculations. To do this, we first compute the temporal derivative of the interpolating statistical pressure,

$$\partial_t A_{N,p,\alpha} = \frac{1}{2}\mathbb{E}\omega_{t,\boldsymbol{x}}(m^p) + \frac{1}{2N\sqrt{tN^{p-1}}}\sqrt{\mathcal{V}}\sum_{\mu}\sum_{i_1<\cdots<i_{p/2}}\mathbb{E}\,J^{\mu}_{i_1,\ldots,i_{p/2}}\omega_{t,\boldsymbol{x}}(\sigma_{i_1}\ldots\sigma_{i_{p/2}}\tau_\mu) - \frac{\mathcal{V}_0}{2}\mathbb{E}\omega_{t,\boldsymbol{x}}(p_{11}).$$

We can apply the Wick–Isserlis theorem on the second contribution to get

$$\begin{aligned}
\partial_t A_{N,p,\alpha} &= \frac{1}{2}\langle m^p\rangle_{t,\boldsymbol{x}} + \frac{\mathcal{V}_0}{2N^p}\frac{p}{2}!\sum_{\mu}\sum_{i_1<\cdots<i_{p/2}}\mathbb{E}\big[\omega_{t,\boldsymbol{x}}(\tau_\mu^2) - \omega^2_{t,\boldsymbol{x}}(\sigma_{i_1}\ldots\sigma_{i_{p/2}}\tau_\mu)\big] - \frac{\mathcal{V}_0}{2}\langle p_{11}\rangle_{t,\boldsymbol{x}}\\
&= \frac{1}{2}\langle m^p\rangle_{t,\boldsymbol{x}} + \frac{\mathcal{V}_0}{2N^p}\left(N^p\langle p_{11}\rangle_{t,\boldsymbol{x}} - N^p\langle q_{12}^{p/2}p_{12}\rangle_{t,\boldsymbol{x}}\right) - \frac{\mathcal{V}_0}{2}\langle p_{11}\rangle_{t,\boldsymbol{x}}\\
&= \frac{1}{2}\langle m^p\rangle_{t,\boldsymbol{x}} - \frac{\mathcal{V}_0}{2}\langle p_{12}q_{12}^{p/2}\rangle_{t,\boldsymbol{x}},
\end{aligned}$$

where we used the fact that $\frac{p}{2}!\sum_{i_1<\cdots<i_{p/2}} \equiv \sum_{i_1,\ldots,i_{p/2}}$ in the thermodynamic limit and the definitions of the overlap order parameters (5.18) and (5.19). Recalling that $S_{N,p,\alpha}(t,\boldsymbol{x}) = 2A_{N,p,\alpha}(t,\boldsymbol{x}) - x$, we finally get the result. □

APPENDIX E: PROOF OF LEMMA 5

Proof. We only prove Eq. (5.24), the other one can be obtained in an analogous way. We will denote for simplicity of notation $\langle\cdot\rangle_{t,\boldsymbol{x}}$ with $\langle\cdot\rangle$. Thus,

$$\begin{aligned}\partial_x\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})\rangle &= \frac{1}{2\sqrt{x}}\sum_{i=1}^N\sum_{a=1}^2 \mathbb{E}_\eta \eta_i \Omega^{(2)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(a)}\Big) - \frac{1}{\sqrt{x}}\sum_{i=1}^N \mathbb{E}\eta_i\Omega^{(3)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\Big)\\ &= \frac{1}{2\sqrt{x}}\sum_{i=1}^N\sum_{a=1}^2 \mathbb{E}_\eta \partial_{\eta_i} \Omega^{(2)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(a)}\Big) - \frac{1}{\sqrt{x}}\sum_{i=1}^N \mathbb{E}\partial_{\eta_i}\Omega^{(3)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\Big),\end{aligned} \tag{E1}$$

where in the last line we used the Wick–Isserlis theorem. Now, it is simple to see that

$$\partial_{\eta_i}\Omega^{(2)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(a)}\Big) = \sqrt{x}\left[\sum_{b=1}^2 \Omega^{(2)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(a)}\sigma_i^{(b)}\Big) - 2\Omega^{(3)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(a)}\sigma_i^{(3)}\Big)\right] \tag{E2}$$

and

$$\partial_{\eta_i}\Omega^{(3)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\Big) = \sqrt{x}\left[\sum_{b=1}^3 \Omega^{(3)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\sigma_i^{(b)}\Big) - 3\Omega^{(4)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\sigma_i^{(4)}\Big)\right]. \tag{E3}$$

By using Eqs. (E2) and (E3) into (E1), we get

$$\begin{aligned}\partial_x\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})\rangle = \frac{1}{2}\sum_{i=1}^N\sum_{a,\;b=1}^2 \mathbb{E}\Omega^{(2)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(a)}\sigma_i^{(b)}\Big) - \sum_{i=1}^N\sum_{b=1}^2 \mathbb{E}\Omega^{(3)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\sigma_i^{(a)}\Big)\\ - \sum_{i=1}^N\sum_{a=1}^3 \mathbb{E}\Omega^{(3)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\sigma_i^{(a)}\Big) + 3\sum_{i=1}^N \mathbb{E}\Omega^{(4)}\Big(O(\boldsymbol{\sigma},\boldsymbol{\tau})\sigma_i^{(3)}\sigma_i^{(4)}\Big).\end{aligned}$$

Recalling that $q_{ab} = \frac{1}{N}\sum_i \sigma_i^{(a)}\sigma_i^{(b)}$,we can write

$$\begin{aligned}\partial_x\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})\rangle &= \frac{N}{2}\sum_{a,b=1}^2 \langle O(\boldsymbol{\sigma},\boldsymbol{\tau})q_{ab}\rangle - \sum_{a=1}^2\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})q_{a3}\rangle - \sum_{a=1}^3\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})q_{3a}\rangle + 3N\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})q_{34}\rangle\\ &= \frac{N}{2}\sum_{a,b=1}^2 \langle O(\boldsymbol{\sigma},\boldsymbol{\tau})q_{ab}\rangle - 2N\sum_{a=1}^2\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})q_{a3}\rangle - N\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})\rangle + 3N\langle O(\boldsymbol{\sigma},\boldsymbol{\tau})q_{34}\rangle,\end{aligned} \tag{E4}$$

thus obtaining Eq. (5.24). □

REFERENCES

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).
[2] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks **61**, 85–117 (2015).
[3] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and beyond: An Introduction to the Replica Method and its Applications* (World Scientific Publishing Company, 1987), Vol. 9.
[4] D. Sherrington and S. Kirkpatrick, "Solvable model of a spin-glass," Phys. Rev. Lett. **35**, 1792 (1975).
[5] G. Parisi, "Toward a mean field theory for spin glasses," Phys. Lett. A **73**, 203–205 (1979).
[6] G. Parisi, "Infinite number of order parameters for spin-glasses," Phys. Rev. Lett. **43**, 1754 (1979).
[7] G. Parisi, "The order parameter for spin glasses: A function on the interval 0-1," J. Phys. A: Math. Gen. **13**, 1101 (1980).
[8] G. Parisi, "A sequence of approximated solutions to the SK model for spin glasses," J. Phys. A: Math. Gen. **13**, L115 (1980).
[9] M. Mézard, G. Parisi, N. Sourlas, G. Toulouse, and M. Virasoro, "Replica symmetry breaking and the nature of the spin glass phase," J. Phys. **45**, 843–854 (1984).
[10] F. Guerra, "Broken replica symmetry bounds in the mean field spin glass model," Commun. Math. Phys. **233**, 1–12 (2003).
[11] S. Ghirlanda and F. Guerra, "General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity," J. Phys. A: Math. Gen. **31**, 9149 (1998).
[12] M. Talagrand, "Replica symmetry breaking and exponential inequalities for the Sherrington-Kirkpatrick model," Ann. Probab. **28**, 1018–1062 (2000).
[13] D. Panchenko, "A connection between the Ghirlanda–Guerra identities and ultrametricity," Ann. Probab. **38**, 327–347 (2010).

[14] D. Panchenko, “Ghirlanda–Guerra identities and ultrametricity: An elementary proof in the discrete case,” C. R. Math. **349**, 813–816 (2011).
[15] D. Panchenko, “The parisi ultrametricity conjecture,” Ann. Math., **177** 383–393 (2013).
[16] D. J. Amit, H. Gutfreund, and H. Sompolinsky, “Storing infinite numbers of patterns in a spin-glass model of neural networks,” Phys. Rev. Lett. **55**, 1530 (1985).
[17] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” Proc. Natl. Acad. Sci. U. S. A. **79**, 2554–2558 (1982).
[18] L. A. Pastur and A. L. Figotin, “Exactly soluble model of a spin glass,” J. Low Temp. Phys. **3**(6), 378–383 (1977).
[19] D. Krotov and J. J. Hopfield, “Dense associative memory for pattern recognition,” in *Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS’16)* (Curran Associates Inc., 2016), pp. 1180–1188.
[20] D. Krotov and J. Hopfield, “Dense associative memory is robust to adversarial inputs,” Neural Comput. **30**, 3151–3167 (2018).
[21] E. Agliari and G. De Marzo, “Tolerance versus synaptic noise in dense associative memories,” Eur. Phys. J. Plus **135**, 1–22 (2020).
[22] P. Baldi and S. S. Venkatesh, “Number of stable points for spin-glasses and neural networks of higher orders,” Phys. Rev. Lett. **58**, 913 (1987).
[23] A. Bovier and B. Niederhauser, “The spin-glass phase-transition in the Hopfield model with p-spin interactions,” Adv. Theoret. Math. Phys. **5**, 1001–1046.
[24] H. Steffan and R. Kühn, “Replica symmetry breaking in attractor neural network models,” Z. Phys. B: Condens. Matter **95**, 249–260 (1994).
[25] T. Tanaka, “Moment problem in replica method,” Interdiscip. Inf. Sci. **13**, 17–23 (2007).
[26] F. Guerra, “Sum rules for the free energy in the mean field spin glass model,” Fields Inst. Commun. **30**(11), 1 (2001).
[27] E. Agliari, A. Barra, R. Burioni, and A. Di Biasio, “Notes on the p-spin glass studied via Hamilton-Jacobi and smooth-cavity techniques,” J. Math. Phys. **53**, 063304 (2012).
[28] A. Barra, G. Dal Ferraro, and D. Tantari, “Mean field spin glasses treated with PDE techniques,” Eur. Phys. J. B **86**, 1–10 (2013).
[29] A. Barra, “The mean field Ising model trough interpolating techniques,” J. Stat. Phys. **132**, 787–809 (2008).
[30] A. Barra, A. D. Biasio, and F. Guerra, “Replica symmetry breaking in mean-field spin glasses through the Hamilton–Jacobi technique,” J. Stat. Mech.: Theory Exp., P09006 (2010).
[31] A. Barra, A. Di Lorenzo, F. Guerra, and A. Moro, “On quantum and relativistic mechanical analogues in mean-field spin models,” Proc. R. Soc. A **470**, 20140589 (2014).
[32] G. De Matteis, F. Giglio, and A. Moro, “Exact equations of state for nematics,” Ann. Phys. **396**, 386–396 (2018).
[33] E. Agliari, D. Migliozzi, and D. Tantari, “Non-convex multi-species Hopfield models,” J. Stat. Phys. **172**, 1247–1269 (2018).
[34] P. Lorenzoni and A. Moro, “Exact analysis of phase transitions in mean-field Potts models,” Phys. Rev. E **100**, 022103 (2019).
[35] E. Agliari, A. Barra, and M. Notarnicola, “The relativistic Hopfield network: Rigorous results,” J. Math. Phys. **60**, 033302 (2019).
[36] A. Fachechi, “PDE/statistical mechanics duality: Relation between Guerra’s interpolated p-spin ferromagnets and the Burgers hierarchy,” J. Stat. Phys. **183**, 1–28 (2021).
[37] E. Agliari, L. Albanese, F. Alemanno, and A. Fachechi, “A transport equation approach for deep neural networks with quenched random weights,” J. Phys. A: Math. Theor. **54**, 505004 (2021).
[38] A. Barra and A. Moro, “Exact solution of the Van der Waals model in the critical region,” Ann. Phys. **359**, 290–299 (2015).
[39] G. De Nittis and A. Moro, “Thermodynamic phase transitions and shock singularities,” Proc. R. Soc. A **468**, 701–719 (2012).
[40] F. Giglio, G. Landolfi, and A. Moro, “Integrable extended Van der Waals model,” Physica D **333**, 293–300 (2016).
[41] A. Moro, “Shock dynamics of phase diagrams,” Ann. Phys. **343**, 49–60 (2014).
[42] B. Derrida, “Random-energy model: An exactly solvable model of disordered systems,” Phys. Rev. B **24**, 2613 (1981).
[43] L. A. Pastur and M. V. Shcherbina, “Absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model,” J. Stat. Phys. **62**, 1–19 (1991).
[44] A. Bovier, “Self-averaging in a class of generalized Hopfield models,” J. Phys. A: Math. Gen. **27**, 7069 (1994).
[45] D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks* (Cambridge University Press, 1989).
[46] E. Agliari, A. Barra, P. Sollich, and L. Zdeborová, “Machine learning and statistical physics: Preface,” J. Phys. A: Math. Theor. **53**, 500401 (2020).
[47] E. Agliari, L. Albanese, A. Barra, and G. Ottaviani, “Replica symmetry breaking in neural networks: A few steps toward rigorous results,” J. Phys. A: Math. Theor. **53**, 415005 (2020).
[48] L. Albanese, F. Alemanno, A. Alessandrelli, and A. Barra, “Replica symmetry breaking in dense neural networks,” J. Stat. Phys **189**(2), 1–41 (2022).
[49] E. Agliari, F. Alemanno, A. Barra, M. Centonze, and A. Fachechi, “Neural networks with a redundant representation: Detecting the undetectable,” Phys. Rev. Lett. **124**, 028301 (2020).
[50] P. Carmona and Y. Hu, “Universality in Sherrington-Kirkpatrick’s spin glass model,” Ann. Inst. Henri Poincare B **42**, 215–222 (2006).
[51] G. Genovese, “Universality in bipartite mean field spin glasses,” J. Math. Phys. **53**, 123304 (2012).
[52] E. Agliari, A. Barra, and B. Tirozzi, “Free energies of Boltzmann machines: Self-averaging, annealed and replica symmetric approximations in the thermodynamic limit,” J. Stat. Mech.: Theory Exp., **2019**, 033301.
[53] E. Gardner, “Multiconnected neural network models,” J. Phys. A: Math. Gen. **20**, 3453 (1987).
[54] E. Agliari, F. Alemanno, A. Barra, and A. Fachechi, “Generalized Guerra’s interpolation schemes for dense associative neural networks,” Neural Networks **128**, 254–267 (2020).
[55] P. J. Olver, “Evolution equations possessing infinitely many symmetries,” J. Math. Phys. **18**, 1212–1215 (1977).
[56] H. Tasso, “Cole’s ansatz and extensions of Burgers’ equation,” Tech. Rep. (Max-Planck-Institut für Plasmaphysik, 1976).
[57] A. Klenke, *Probability Theory: A Comprehensive Course* (Springer Science & Business Media, 2013).